Acerca de este libro

Esta es una copia digital de un libro que, durante generaciones, se ha conservado en las estanterías de una biblioteca, hasta que Google ha decidido escanearlo como parte de un proyecto que pretende que sea posible descubrir en línea libros de todo el mundo.

Ha sobrevivido tantos años como para que los derechos de autor hayan expirado y el libro pase a ser de dominio público. El que un libro sea de dominio público significa que nunca ha estado protegido por derechos de autor, o bien que el período legal de estos derechos ya ha expirado. Es posible que una misma obra sea de dominio público en unos países y, sin embargo, no lo sea en otros. Los libros de dominio público son nuestras puertas hacia el pasado, suponen un patrimonio histórico, cultural y de conocimientos que, a menudo, resulta difícil de descubrir.

Todas las anotaciones, marcas y otras señales en los márgenes que estén presentes en el volumen original aparecerán también en este archivo como testimonio del largo viaje que el libro ha recorrido desde el editor hasta la biblioteca y, finalmente, hasta usted.

Normas de uso

Google se enorgullece de poder colaborar con distintas bibliotecas para digitalizar los materiales de dominio público a fin de hacerlos accesibles a todo el mundo. Los libros de dominio público son patrimonio de todos, nosotros somos sus humildes guardianes. No obstante, se trata de un trabajo caro. Por este motivo, y para poder ofrecer este recurso, hemos tomado medidas para evitar que se produzca un abuso por parte de terceros con fines comerciales, y hemos incluido restricciones técnicas sobre las solicitudes automatizadas.

Asimismo, le pedimos que:

+ *Haga un uso exclusivamente no comercial de estos archivos* Hemos diseñado la Búsqueda de libros de Google para el uso de particulares; como tal, le pedimos que utilice estos archivos con fines personales, y no comerciales.

+ *No envíe solicitudes automatizadas* Por favor, no envíe solicitudes automatizadas de ningún tipo al sistema de Google. Si está llevando a cabo una investigación sobre traducción automática, reconocimiento óptico de caracteres u otros campos para los que resulte útil disfrutar de acceso a una gran cantidad de texto, por favor, envíenos un mensaje. Fomentamos el uso de materiales de dominio público con estos propósitos y seguro que podremos ayudarle.

+ *Conserve la atribución* La filigrana de Google que verá en todos los archivos es fundamental para informar a los usuarios sobre este proyecto y ayudarles a encontrar materiales adicionales en la Búsqueda de libros de Google. Por favor, no la elimine.

+ *Manténgase siempre dentro de la legalidad* Sea cual sea el uso que haga de estos materiales, recuerde que es responsable de asegurarse de que todo lo que hace es legal. No dé por sentado que, por el hecho de que una obra se considere de dominio público para los usuarios de los Estados Unidos, lo será también para los usuarios de otros países. La legislación sobre derechos de autor varía de un país a otro, y no podemos facilitar información sobre si está permitido un uso específico de algún libro. Por favor, no suponga que la aparición de un libro en nuestro programa significa que se puede utilizar de igual manera en todo el mundo. La responsabilidad ante la infracción de los derechos de autor puede ser muy grave.

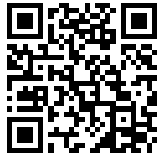
Acerca de la Búsqueda de libros de Google

El objetivo de Google consiste en organizar información procedente de todo el mundo y hacerla accesible y útil de forma universal. El programa de Búsqueda de libros de Google ayuda a los lectores a descubrir los libros de todo el mundo a la vez que ayuda a autores y editores a llegar a nuevas audiencias. Podrá realizar búsquedas en el texto completo de este libro en la web, en la página `http://books.google.com`

This is a reproduction of a library book that was digitized by Google as part of an ongoing effort to preserve the information in books and make it universally accessible.

Google™ books

https://books.google.com



Digitized by Google

STANFORD·VNIVERSITY·LIBRARY

Digitized by Google

Digitized by Google

Digitized by Google

PROCEEDINGS

OF THE

CAMBRIDGE PHILOSOPHICAL SOCIETY

VOLUME XXIII

Digitized by Google

CAMBRIDGE: PRINTED BY
W. LEWIS, M.A.
AT THE UNIVERSITY PRESS

Digitized by Google

PROCEEDINGS

OF THE

CAMBRIDGE PHILOSOPHICAL SOCIETY

VOLUME XXIII

26 OCTOBER 1925—25 JULY 1927

CAMBRIDGE

AT THE UNIVERSITY PRESS

AND SOLD BY

DEIGHTON, BELL & CO., LTD. AND BOWES & BOWES, CAMBRIDGE

CAMBRIDGE UNIVERSITY PRESS, FETTER LANE, LONDON, E.C. 4

1927

Digitized by Google

STANFORD·VNIVERSITY·LIBRARY

Digitized by Google

Digitized by Google

Digitized by Google

Digitized by Google

Digitized by Google

PROCEEDINGS

OF THE

CAMBRIDGE PHILOSOPHICAL SOCIETY

VOLUME XXIII

Digitized by Google

CAMBRIDGE: PRINTED BY
W. LEWIS, M.A.
AT THE UNIVERSITY PRESS

Digitized by Google

PROCEEDINGS

OF THE

CAMBRIDGE PHILOSOPHICAL SOCIETY

VOLUME XXIII

26 OCTOBER 1925—25 JULY 1927

CAMBRIDGE
AT THE UNIVERSITY PRESS
AND SOLD BY
DEIGHTON, BELL & CO., LTD. AND BOWES & BOWES, CAMBRIDGE
CAMBRIDGE UNIVERSITY PRESS, FETTER LANE, LONDON, E.C. 4
1927

Digitized by Google

360911

CAMROSE LIBRARY

PRINTED IN GREAT BRITAIN

Digitized by Google

CONTENTS

VOL. XXIII

PAGE

The stability of rotating masses of liquid; note to a former paper. By Professor H. F. BAKER 1

On the hypothesis of the obtuse angle. By Professor M. J. M. HILL, Peterhouse. (With Twenty-Six Text-figures) 2

On the substitution of Wallis's postulate of similarity for Euclid's postulate of parallels. Addendum. By Professor M. J. M. HILL, Peterhouse. (With Three Text-figures) 19

The reciprocation of one quadric into another. By Professor H. F. BAKER 22

Note on the pedal locus. By Mr J. P. GABBATT, Peterhouse . . . 28

On the series of forms of Jacobi's rotating liquid ellipsoid. By W. F. SEDGWICK, M.A., Trinity College 33

Periodic solutions of a linear differential equation of the second order with periodic coefficients. By Mr E. L. INCE, Trinity College . . 44

The second solution of the Mathieu equation. By Mr E. L. INCE, Trinity College 47

On a family of constructs in higher space. By C. G. F. JAMES, B.A., Trinity College. (With Two Text-figures) 50

The equianharmonic and harmonic envelopes of a quadrinodal cubic surface. By G. TIMMS, St John's College. (Communicated by Mr W. P. MILNE) 68

The adiabatic hypothesis for magnetic fields. By P. A. M. DIRAC, St John's College. (Communicated by Mr R. H. FOWLER) . . 69

A photographic method of determining the mobility of recoil atoms. By G. H. BRIGGS. (Communicated by Dr J. CHADWICK.) (With One Text-figure and Plate I) 73

The quenching of resonance radiation and the breadths of absorption-lines. By R. W. DITCHBURN, B.A., Trinity College 78

Notes on the electrical counter. By K. G. EMELÉUS, B.A., St John's College. (Communicated by Mr E. V. APPLETON.) (With Three Text-figures) 85

On chains of two-two relations and the theory of elliptic functions. By Professor H. F. BAKER 92

On a group of 1440 *birational transformations of four variables that arises in considering the projective equivalence of double sixes.* By Mr W. BURNSIDE, Honorary Fellow of Pembroke College . . 103

The general (m, n) *correspondence.* By R. VAIDYANATHASWAMY. (Communicated by Mr H. W. TURNBULL) 109

Digitized by Google

PAGE

The stability of a strut under thrust, when buckling is resisted by a force proportional to the displacement. By S. GOLDSTEIN, B.A., St John's College. (With One Text-figure) 120

Insular gravity and oceanic isostasy. By Sir JOSEPH LARMOR, Lucasian Professor. 130

A contribution to the theory of ferromagnetism. By G. S. MAHAJANI, B.A., St John's College 136

Over-tension in a condenser battery during a sudden discharge. By P. KAPITZA, Ph.D., Fellow of Trinity College, Assistant Director of Magnetic Research at the Cavendish Laboratory, Clerk Maxwell Student of Cambridge University. (With One Text-figure and Plate II). 144

The period of decay of radium B and radium C. By P. BRACELIN. (Communicated by Dr J. CHADWICK) 150

On the diurnal variation of ultra-short wave wireless transmission. By Mr E. V. APPLETON, St John's College. (With Three Text-figures) 155

The influence of an illuminated mercury surface on the Franck-Cario reactions. By HERBERT SIM HIRST, Trinity College. (Communicated by Mr E. K. RIDEAL.) (With One Text-figure) . . . 162

The behaviour of the centres of activity of saturated surfaces during the initial stages of unimolecular reactions. By F. H. CONSTABLE, B.A., Fellow of St John's College. (With Two Text-figures) . . . 172

Metallic complexes with the aliphatic poly-amines. By Dr F. G. MANN and Prof. Sir WILLIAM J. POPE. (With One Text-figure) . . 183

Oxidations on charcoal. By Miss W. M. WRIGHT. (Communicated by Mr E. K. RIDEAL) 187

The problem of relativity in reference to several bodies. By Mr R. HARGREAVES, St John's College 191

Relativity in connexion with axial rotation. By Mr R. HARGREAVES, St John's College 198

Sur la non-existence simultanée de deux fonctions de Mathieu. Par Ž. MARKOVIĆ. (Communicated by Mr E. L. INCE) 203

On the higher singularities of plane algebraic curves. By Miss C. A. SCOTT. (Communicated by Mr F. P. WHITE.) (With One Text-figure) 206

The (2, 1) *correspondence.* By R. VAIDYANATHASWAMY. (Communicated by Mr H. W. TURNBULL) 233

Integral invariants of the affine field. By Mr M. H. A. NEWMAN, St John's College 262

An approximation to the motion of two rotating electrical doublets in a plane. By P. A. TAYLOR, B.A., Emmanuel College. (With Two Text-figures) 269

A theory of the action of X-rays on living cells. By J. A. CROWTHER, Sc.D., St John's College. (With One Text-figure) 284

Digitized by Google

PAGE

On the attenuation of wireless signals in short distance overland transmission. By J. A. RATCLIFFE, B.A. and M. A. F. BARNETT. (Communicated by Mr E. V. APPLETON.) (With Five Text-figures) 288

Some relations between the optical spectra of different atoms of the same electronic structure. II. *Aluminium-like and copper-like atoms.* By Mr D. R. HARTREE, Fellow of St John's College . . . 304

On the quantum theory of vibration-rotation bands. By J. R. OPPENHEIMER. (Communicated by Mr R. H. FOWLER) 327

The elliptic products of Jacobi and the theory of linear congruences. By Major P. A. MACMAHON 337

On Mellin's inversion formula. By Mr J. C. BURKILL, Trinity College . 356

Note on the extension to higher space of a theorem of Wallace. By Mr J. P. GABBATT, Peterhouse 361

A mathematical theory of natural and artificial selection. Part III. By Mr J. B. S. HALDANE, Trinity College 363

The summation of a Fourier integral of finite type. By Mr S. POLLARD, Trinity College 373

On the descriptive form of Taylor's theorem. By Mr S. POLLARD, Trinity College 383

The correspondence between lines in threefold space and points of a quadric fourfold in fivefold space, established by a geometrical construction. By Mr T. L. WREN, St John's College . . . 386

On some resistance properties of a certain network containing inductances and capacities and their analogies in a vibrating mechanical system. By Mr E. B. MOULLIN, Downing College. (Communicated by Dr G. F. C. SEARLE.) (With Seven Text-figures) 391

The polarisabilities of atomic cores. By BERTHA SWIRLES, Yarrow Research Student, Girton College. (Communicated by Mr R. H. FOWLER.) (With Two Text-figures) 403

On quantum algebra. By Dr P. A. M. DIRAC, St John's College. (Communicated by Mr R. H. FOWLER) 412

An optical illusion due to contrast. By Mr G. P. THOMSON, Corpus Christi College. (With One Text-figure and Plate III) . . . 419

On the quantum theory of the problem of the two bodies. (Preliminary Communication.) By Mr J. R. OPPENHEIMER. (Communicated by Mr R. H. FOWLER) 422

On the stability of copper catalysts produced by thermal decomposition. By Dr F. H. CONSTABLE, Fellow of St John's College . . . 432

Note on the velocity of gas-reactions. By Mr J. A. CHRISTIANSEN, of the University of Copenhagen. (Communicated by Mr R. H. FOWLER) 438

Improvements in the apparatus for measuring the velocity of very rapid chemical reactions. II. By Dr H. HARTRIDGE, Fellow of King's College, Cambridge, and Dr F. J. W. ROUGHTON, Fellow of Trinity College, Cambridge. (With Five Text-figures) 450

Digitized by Google

PAG

Note on the Petzval optical condition. By Mr G. C. STEWARD, Gonville and Caius College. (With One Text-figure) 46

Maxwell's law, and the absorption and emission of radiation. By Mr E. A. MILNE, Trinity College 46

On compressional waves in two superposed layers. By Mr HAROLD JEFFREYS, St John's College 47

Proceedings at the meetings held during the session 1925—1926 . . 48

Statement of Accounts of the Cambridge Philosophical Society, from January 1, 1925, to December 31, 1925 49

Geodetic and dynamical principles, a comparison and connexion. By Mr R. HARGREAVES, St John's College 49

On the figure of Pappus' theorem. By Mr C. V. HANUMANTA RAO, Trinity College. (Communicated by Professor H. F. BAKER) . 49

On a group of order 25920 *and the projective transformations of a cubic surface.* By Mr W. BURNSIDE, Pembroke College 49

The Compton effect in wave mechanics. By Dr P. A. M. DIRAC, St John's College 500

On the polarisation of mercury lines emitted from a discharge tube in a magnetic field. By Dr H. W. B. SKINNER, Trinity College. (With Three Text-figures) 508

Note on the effect of alpha particles on paraffin. By WILLIAM T. RICHARDS, Ph.D. (Communicated by Prof. Sir E. RUTHERFORD.) (With Two Text-figures) 516

The application of the method of the magnetic spectrum to the study of secondary electronic emission. By C. F. SHARMAN, B.A., King's College. (Communicated by Prof. Sir E. RUTHERFORD.) (With Four Text-figures) 523

Some investigations of gas discharges by means of an exploring electrode. By Dr K. G. EMELÉUS, St John's College. (Communicated by Mr E. V. APPLETON.) (With Three Text-figures) 531

The calculation of atomic fields. By L. H. THOMAS, B.A., Trinity College. (With One Text-figure) 542

The tensile deformation of large aluminium crystals at crystal boundaries. By R. L. ASTON, M.Sc., Barker Graduate Scholar of the University of Sydney, Dominion Exhibitioner of Trinity College. (Communicated by Mr G. I. TAYLOR.) (With Nine Text-figures and Plate IV) 549

Some developments in the X-ray analysis of single crystals. By R. L. ASTON, M.Sc., Barker Graduate Scholar of the University of Sydney, Dominion Exhibitioner of Trinity College. (Communicated by Mr G. I. TAYLOR.) (With Eleven Text-figures) 561

Contribution to the theory of the diffusion pump. By L. WERTENSTEIN, D.Sc., Professor of Radiology at the Free-University, Warsaw. (Communicated by Professor Sir E. RUTHERFORD.) (With One Text-figure) 578

Digitized by Google

PAGE

On the addition of the primary aberrations. By Mr G. C. Steward, Gonville and Caius College 584

The absorption spectra of "saturated" and "unsaturated" organic substances. By Mr J. E. Purvis. (With One Text-figure) . . 588

Surface adsorption and the velocity of chemical action at gas-solid interfaces. By Dr F. Hurn Constable, Fellow of St John's College. (With Three Text-figures) 593

A mathematical theory of natural and artificial selection. Part IV. By Mr J. B. S. Haldane, Trinity College 607

What determines the resistance and the tilt of an aeroplane? By Sir Joseph Larmor, Lucasian Professor. (With One Text-figure) . 617

The pedal (3, 2) *correspondence.* By R. Vaidyanathaswamy. (Communicated by Mr H. W. Turnbull) 631

Lines and planes in a metrical space. By W. Hunter, B.A., Trinity College. (Communicated by Mr S. Pollard.) (With Four Text-figures) 653

Note on Ramanujan's arithmetical function $\tau(n)$. By Mr G. H. Hardy, Trinity College 675

Notes on the theory of series (III): *On the summability of the Fourier series of a nearly continuous function.* By Mr G. H. Hardy and Mr J. E. Littlewood, Trinity College 681

The evaluation of Gibbs' phase-integral for imperfect gases. By H. D. Ursell, Trinity College. (Communicated by Mr R. H. Fowler). 685

The limits of classical scattering. By P. M. S. Blackett, M.A., Fellow of King's College 698

On Herschel's condition and the optical cosine law. By Mr G. C. Steward, Gonville and Caius College. (With One Text-figure) . 703

The effect of the orbital velocity of the electrons in heavy atoms on their stopping of α-particles. By L. H. Thomas, B.A., Trinity College. (With One Text-figure) 713

The absolute intensities of the γ-rays of radium B and radium C. By C. D. Ellis, Ph.D., Trinity College, and W. A. Wooster, B.A., Peterhouse 717

An experiment on the stability of superposed streams of fluid. By Mr G. I. Taylor, Trinity College. (With One Text-figure) . . 730

The stopping power of hydrogen atoms for α-particles according to the new quantum theory. By J. A. Gaunt, B.A., Trinity College. (Communicated by Mr R. H. Fowler.) (With One Text-figure) . 732

A case of distinction between Fourier integrals and Fourier series. By Margaret Eleanor Grimshaw, Newnham College. (Communicated by Mr S. Pollard) 755

Wave propagation in strings with continuous and concentrated loads. By Mr Harold Jeffreys, St John's College 768

On the fifth book of Euclid's Elements. (Second Addendum to Fifth Paper.) Proof of Euclid v, 19. By Dr M. J. M. Hill, Peterhouse 779

Digitized by Google

PAGE

Some measurements on the absorption of X-rays of long wave-length. By L. H. MARTIN, M.Sc., Trinity College. (Communicated by Prof. Sir E. RUTHERFORD.) (With Two Text-figures) . . . 783

The total energy of binding of a heavy atom. By Mr E. A. MILNE, Trinity College 794

The Hall effect in single crystals of iron. By Dr W. L. WEBSTER, Trinity College. (Communicated by Dr KAPITZA.) (With One Text-figure) 800

Excitation by high velocity electrons. By Mr E. G. DYMOND, Fellow of St John's College. (With Two Text-figures) 804

A valve amplifier for ionisation currents. By C. E. WYNN-WILLIAMS, Trinity College. (Communicated by Prof. Sir E. RUTHERFORD.) (With Four Text-figures) 811

The production of characteristic X-rays by electronic impact. By L. H. THOMAS, B.A., Trinity College. (With One Text-figure) . 829

On the effect of continued small additions of poisonous substances on the velocity of gaseous catalytic reaction in closed vessels. By Dr F. HURN CONSTABLE, Fellow of St John's College . . . 832

A mathematical theory of natural and artificial selection, Part V: Selection and mutation. By Mr J. B. S. HALDANE, Trinity College 838

A theorem of Dr P. Zeeman. By Mr J. H. GRACE, Peterhouse . . 845

The pedal planes of a tetrahedron. By Mr J. H. GRACE, Peterhouse 853

An illustration of the space representation of circles. By Mr J. H. GRACE, Peterhouse 859

On the generalizations of the theorems of Parseval and Riesz-Fischer. By Mr S. POLLARD, Trinity College 865

Summation of the integral conjugate to the Fourier integral of finite type. By MARGARET ELEANOR GRIMSHAW, Newnham College. (Communicated by Mr S. POLLARD) 871

Simplexes and other configurations upon a rational normal curve. By Mr F. P. WHITE, St John's College 882

The specific heat of carbon dioxide and the form of the CO_2 molecule. By W. H. M^{c}CREA, B.A., Trinity College. (Communicated by Mr R. H. FOWLER.) (With Two Text-figures) 890

Eddy currents in thin circular cylinders of uniform conductivity, due to periodically changing magnetic fields, in two dimensions. By Dr F. W. CARTER, St John's College. (Communicated by Mr F. P. WHITE) 901

The relation between refractive index and density. By D. BURNETT, B.A., Clare College 907

On the distribution of the error of an interpolated value, and on the construction of tables. By Dr R. A. FISHER, Gonville and Caius College, and J. WISHART 912

Digitized by Google

PAGE

A differential retarding potential method for the study of the energy distribution of slow electron emissions. By C. F. SHARMAN, B.A., King's College. (Communicated by Prof. Sir E. RUTHERFORD.) (With Five Text-figures) 922

The emission of radiation by a quadripole electric moment on the quantum mechanics. By J. A. GAUNT, B.A., Trinity College, and W. H. M^{c}CREA, B.A., Trinity College 930

The amount of energy emitted in the γ-ray form by radium E. By G. H. ASTON, B.A., Trinity Hall. (Communicated by Prof. Sir E. RUTHERFORD.) (With One Text-figure) 935

The specific heat of water vapour and the theory of the dissociation of water vapour at high temperatures. By W. H. M^{c}CREA, B.A., Trinity College. (With One Text-figure) 942

A case of double reflexion. By E. T. S. APPLEYARD, B.A., King's College, and Dr H. W. B. SKINNER, Trinity College. (With One Text-figure) 951

A convenient method of distillation of the alkali metals. By R. J. CLARK, Lecturer in Physics and Carnegie Teaching Fellow in the University of Edinburgh. (Communicated by Prof. Sir E. RUTHERFORD.) (With One Text-figure) 953

A rapid mercury still. By R. J. CLARK, Lecturer in Physics and Carnegie Teaching Fellow in the University of Edinburgh. (Communicated by Prof. Sir E. RUTHERFORD.) (With Three Text-figures) 956

Notes on spectrophotometry. By R. W. DITCHBURN, B.A., Trinity College, Isaac Newton Student. (With Seven Text-figures) . . . 959

The absorption and reduction in velocity of β-rays on their passage through matter. By Dr E. MADGWICK, M.C., Emmanuel College. (Communicated by Prof. Sir E. RUTHERFORD.) (With Nine Text-figures) 970

The β-ray spectrum of Ra E. By Dr E. MADGWICK, M.C., Emmanuel College. (Communicated by Prof. Sir E. RUTHERFORD.) (With One Text-figure) 982

Proceedings at the meetings held during the session 1926—1927 . . 985

Statement of Accounts of the Cambridge Philosophical Society, from January 1, 1926, to December 31, 1926 994

Index to Vol. XXIII 996

PLATES

Plate I. To illustrate Mr Briggs's paper 73

Plate II. To illustrate Dr Kapitza's paper 144

Plate III. To illustrate Mr Thomson's paper 419

Plate IV. To illustrate Mr Aston's paper 549

Digitized by Google

Digitized by Google

PROCEEDINGS

OF THE

Cambridge Philosophical Society.

The stability of rotating masses of liquid; note to a former paper.

By Professor H. F. BAKER.

[*Received* and *read* 4 May 1925.]

In these *Proceedings*, Vol. xx (1920), pp. 198–204, the writer ventured (*l.c.* p. 203) on some remarks as to the reason why Sir G. Darwin and Mr Jeans had obtained discordant results for the stability of the pear-shaped figure of equilibrium, suggesting that this was due to a different mode of expansion of the functions involved. Sir G. Darwin used an infinite series of Lamé functions: Mr Jeans' method was equivalent to using the early parts from an expansion, of which every term, when expressed as an integral series of Lamé functions, would be an infinite series. At that time there was difficulty in obtaining Liapounoff's* papers; since then, by the kindness of M. Belopolsky, of Pulkova, the whole of the four parts of Liapounoff's publication "Sur les figures d'équilibre peu différentes des ellipsoïdes d'une masse liquide homogène douée d'un mouvement de rotation," in all over 750 large folio pages, have become available; and these are now in the market. It is particularly interesting to see that the fourth part (1914) is devoted precisely to that change in the method of development which would arise in passing from Sir G. Darwin's expansion to the other expansion referred to above. And it is only by this change that Liapounoff is able to give the general proof of a form of expression of his results—in terms of polynomials and not infinite series—upon which his theorem of instability is made to depend. The careful consideration of the convergence of his expansions, which adds so greatly to the length of Liapounoff's papers, supplies materials for the proof that the expansion used in Mr Jeans' paper can be placed on a sure foundation, while Sir G. Darwin's expansion requires an estimation of the remainder.

* An obituary notice of Liapounoff, in Russian, was written by Stekloff, *Bulletin de l'Académie des Sciences de Russie*, 1919, pp. 367–388.

Digitized by Google

On the hypothesis of the obtuse angle. By Professor M. J. M. HILL, Peterhouse.

[*Received* 3 April, *read* 4 May, 1925.]

INTRODUCTION.

1. Engel and Stäckel in their *Theorie der Parallellinien von Euklid bis auf Gauss** (1895) call attention to the fact that Saccheri and Lambert have used in their argument the property of the triangle that the external angle is greater than either of the interior and opposite angles (Euc. I. 16), which is not true for triangles of all magnitudes when the hypothesis of the obtuse angle holds good. They do not show how *starting from the point of view of Saccheri and Lambert* the proofs should be amended so that they would be valid in the Elliptic Plane. Bonola (*Non-Euclidean Geometry*, translated by Carslaw, 1912, p. 31) in his account of the subject also uses Euc. I. 16. It is the object of this paper to obtain the main results proved by Saccheri and Lambert when the hypothesis of the obtuse angle holds *without using* Euc. I. 16. It is not suggested that this is a convenient method of procedure but it may perhaps have some historical interest. The key to the procedure here adopted is found in the reversal of the Euclidean order in a certain group of propositions, viz. I. 16, 18, 19, 20.

Euclid, after proving I. 16, deduces I. 17 (any two angles of a triangle are together less than two right angles). Then from I. 16 he deduces I. 18 (the greater side of any triangle is opposite to the greater angle), and from I. 18 its converse I. 19, from which he deduces I. 20 (the sum of any two sides of a triangle is greater than the third side).

Now propositions I. 18, 19 and 20 are true for triangles of all magnitudes in the Elliptic Plane, although I. 16 is not. It is, therefore, desirable to obtain proofs of propositions I. 18, 19 and 20, which do not depend on I. 16.

After these have been given a proof of the following proposition, which gives as much of the information furnished by I. 16 as is required, will be obtained.

Let ABC be a triangle, right-angled at A, but acute-angled at B and C†.

Then if a point P move from B to A, CP steadily decreases, whilst the angle CPA steadily increases.

2. The following propositions in the First Book of Euclid will be assumed: 1–15, 23 and the first part of 26.

* Pp. 52, 144 and elsewhere.

† This is the only kind of right-angled triangle which Saccheri considered.

Digitized by Google

Further, after proofs of Euc. I. 18 and 19, which are independent of I. 16, have been given, it will be possible to use also Euclid's proofs of I. 24 and 25.

In order to forestall criticisms that might be raised from the point of view of Elliptic Geometry, frequent references are made in this paper to the known properties of right-angled triangles in the Elliptic Plane for the purpose of showing that the conclusions in the paper which might have been reached by Saccheri and Lambert are not in conflict with the theorems of Elliptic Geometry. These are in square brackets. If Saccheri had proved his propositions in the manner here set forth, the possibility of the objections noted in square brackets would not have occurred to him, but the proofs would not have been open to such a criticism as that of his use of Euc. I. 16.

3. In view of the alterations here made in Euclid's order of proof it is necessary to give the proofs of some properties of right-angled triangles, which might be more simply proved if Euclid's order could be followed when discussing the hypothesis of the obtuse angle.

First Group of Theorems*.

Theorem I.

4. Let OX, OY be two rays bounding an acute angle.

Let P be any point on OY.

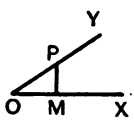
Then the perpendicular drawn from P on to the straight line, of which the ray OX is a part, falls on the ray OX. (The proof is left to the reader.)

Theorem II.

5. There cannot be two distinct lines which are both perpendiculars from P on OX.

Saccheri and Lambert would have dismissed the possibility of two such perpendiculars as being in conflict with Euc. I. 17.

[In Elliptic Geometry, if PM, PN are both perpendicular to OX *and are not situated on the same straight line*, then it is known that every straight line through P is perpendicular to OX.



If then the angle MOP in the figure be acute, it is impossible to draw two distinct lines from P perpendicular to OX.

This leaves open the possibility that MP produced through P may meet OX at right angles.

* The proofs of the theorems of this group are valid whether Euclid's Postulate of Parallels holds good or not.

Digitized by Google

Further it is known in Elliptic Geometry that if MOP is acute, MP is less than the polar distance which we shall call L.

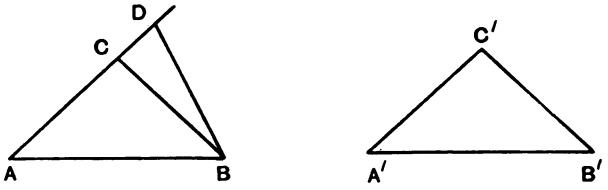
Hence in discussing any right-angled triangle the side opposite to an acute angle is less than L, and if it be known that one of the sides containing the right angle is less than L, then the angle opposite to it is acute.

To represent these facts, which were not known to Saccheri, we shall place in this investigation a limitation on the lengths of certain lines.]

Theorem III.

6. If two right-angled triangles have their hypotenuses equal and an acute angle in each equal, then the triangles are congruent.

In the triangles ABC, $A'B'C'$ let the angles at C and C' be right angles, let AB be equal to $A'B'$, and let the angles BAC, $B'A'C'$ be equal to one another and let both of them be acute.



On AC take AD equal to $A'C'$.

If possible let D be distinct from C. Join BD.

Then the triangles ABD, $A'B'C'$ are congruent. Therefore the angle ADB is equal to $A'C'B'$ and is therefore a right angle.

Therefore BC, BD, which are two distinct lines, are both perpendicular to ACD, whilst the angle BAC is acute, and this is impossible by Art. 5. Hence D coincides with C and therefore the triangles ABC, $A'B'C'$ are congruent*.

Theorem IV.

7. If two right-angled triangles have their hypotenuses equal and a side in each equal, then the triangles are congruent.

In the triangles ABC, $A'B'C'$ let the angles at C and C' be right angles, let AB be equal to $A'B'$ and let the sides AC, $A'C'$ be equal to one another.

Construct the angle BAD equal to the angle $B'A'C'$ on the opposite side of AB to C, and make AD equal to $A'C'$. Join BD. Then the triangles BAD, $B'A'C'$ are congruent.

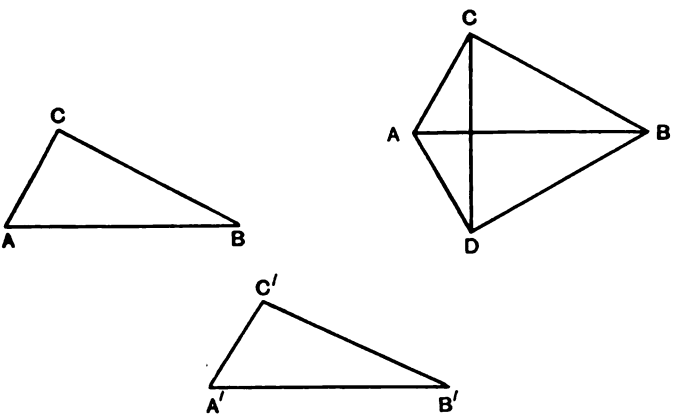
* A similar demonstration would hold good in Elliptic Geometry if the angles BAC, $B'A'C'$ were equal and obtuse.

Digitized by Google

Now AD is equal to $A'C'$ and therefore equal to AC. BD is equal to $B'C'$.

The angle ADB is equal to $A'C'B'$ and is therefore a right angle, and so is equal to ACB. Join CD. Then, since AD is equal to AC, it follows that the angles ACD and ADC are equal and therefore the angles BCD and BDC are equal, and therefore BC is equal to BD and therefore equal to $B'C'$. But AC is equal to $A'C'$ and the angles ACB, $A'C'B'$ are equal; therefore the triangles ABC, $A'B'C'$ are congruent.

[In Elliptic Geometry there would be different figures according as either of the lengths AB, AC were or were not less than L.]

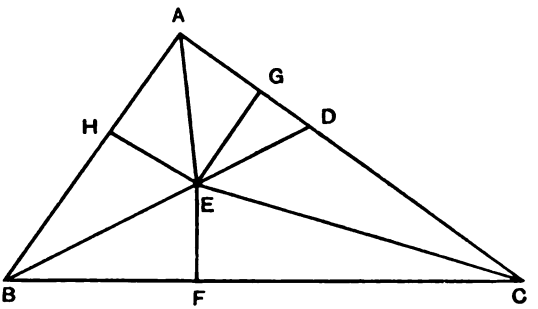


Theorem V.

8. To prove that in a triangle, each of whose angles is less than two right angles, the sum of any two sides is greater than the third side (Euc. I. 20).

Let ABC be such a triangle.

Bisect the angle ABC by BD meeting AC in a point D, necessarily lying between A, C. Then BD is wholly inside the triangle.



Digitized by Google

Bisect the angle BCD by a straight line CE meeting BD in E necessarily between B and D, so that E is a point inside the triangle.

Draw EF, EG, EH perpendicular to BC, CA, AB respectively.

Since ABC is less than two right angles, the angle EBC is acute, and therefore BF is in the direction of the ray BC (Art. 4).

In like manner ECF is acute and therefore F is on the ray CB. So F is on the same side of B as C, and also on the same side of C as B. Therefore F lies between B and C.

Then the triangles EBF, EBH are congruent by Art. 6.

Therefore EF, BF are equal to EH, BH respectively.

In like manner EF, CF are equal to EG, CG respectively. Hence EG is equal to EH. Join EA.

Then in the triangles EAG, EAH there is a right angle in each, the hypotenuses are equal and there is a side in each equal. Therefore the triangles are congruent (Art. 7). Therefore the angles EAG, EAH are equal. Now E is inside the triangle. Therefore EA is the internal bisector of the angle BAC.

Since EAG, EAH are acute angles it follows, just as in the proof that the point F lies between B and C, that G lies between C and A and H lies between B and A.

But BF is equal to BH and is therefore less than BA.

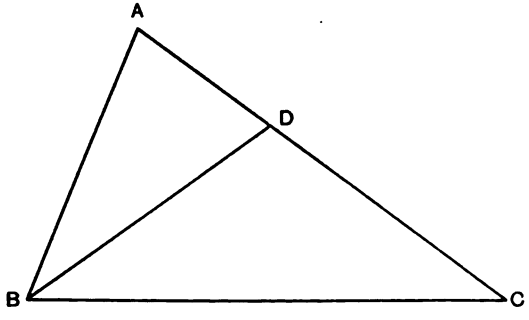
Also CF is equal to CG and is therefore less than CA.

Hence the sum of BF and CF is less than the sum of BA and CA.

Hence BC is less than the sum of BA and CA.

Theorem VI.

9. The angles of the triangle ABC being each less than two right angles and the angle ABC being greater than the angle ACB, it is required to prove that the side AC is greater than the side AB (Euc. I. 19)*.



* The only other proof of this theorem, known to me, which is valid for triangles of any magnitude in the Elliptic Plane, will be found in Coolidge's *Non-Euclidean Geometry*, p. 35, Theorem 28. It involves an axiom of continuity, an infinite limiting process, and a reductio ad absurdum.

Digitized by Google

Through B draw a straight line BD making the angle CBD equal to the angle ACB; then the straight line BD lies between BC and BA and therefore meets AC in some point D between A and C.

Also BD is equal to DC.

Then CA is equal to $CD + DA$ which is equal to $BD + DA$, which is greater than BA by Art. 8.

Theorem VII.

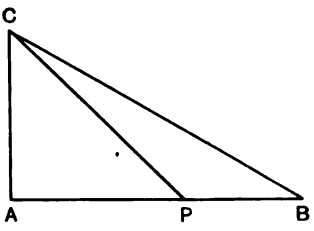
10. The angles of the triangle ABC being each less than two right angles and the side AC being greater than the side AB, it is required to prove that the angle ABC is greater than the angle ACB (Euc. I. 18).

This can be deduced from the preceding theorem by a reductio ad absurdum, or can be proved directly by drawing a perpendicular through its middle point to the side BC of the figure of Art. 9, proving that it must meet AC, the longer of the two sides AB, AC in some point D, and then joining B to D. Then it follows that the angle ABC is greater than the angle ACB.

Theorem VIII.

11. The triangle ABC has a right angle at A and acute angles at B and C*.

To prove that if a point P describe BA from B to A, CP will decrease steadily from CB to CA, whilst the angle CPA will increase steadily from CBA to CAB.



Since CBA is acute it is less than CAB.

Therefore CB is greater than CA. Now produce CA to D so that CA is equal to AD.

Join DP, DB. Then CP and DP are equal, as also are the angles CPA and DPA; CB and DB are equal, as also are the angles CBA, DBA.

In the right-angled triangle CAP the angle CPA [being opposite to CA, which is less than L] is less than a right angle. Therefore the angle CPB is obtuse, but the angle CBP is acute; therefore the angle CPB is greater than the angle CBP.

Therefore CB is greater than CP.

Also CPA is a right-angled triangle and the angles at C, P are each acute.

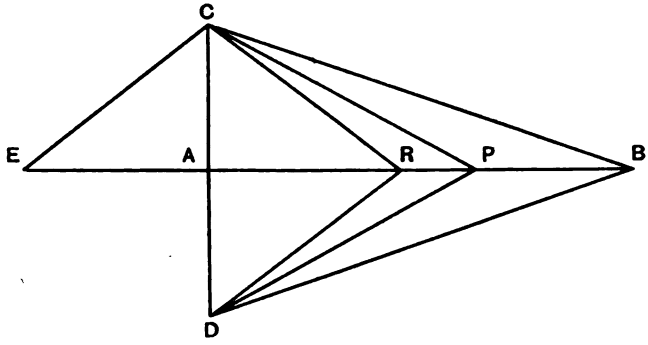
* [In Elliptic Geometry this involves the sides of this triangle, each being less than L.]

Digitized by Google

Now take R on AP between A and P, then it can be proved in like manner that CP is greater than CR and that the angle CRA is acute.

This proves that as P goes from B to A, CP steadily decreases.

It remains to prove that the angle CRA is greater than the angle CPA. Produce RA to E making AE equal to AR. Join EC.



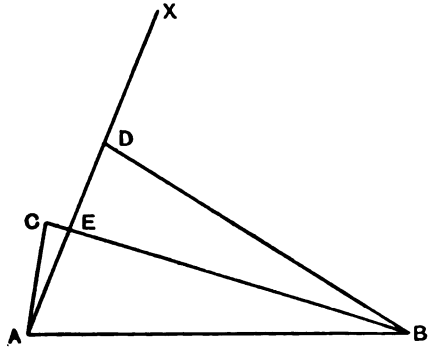
The triangles CAE, CAR are congruent.

Therefore CE is equal to CR, which is less than CP, therefore the angle CEA is greater than the angle CPA. But the angles CEA, CRA are equal; therefore the angle CRA is greater than the angle CPA.

Theorem IX.

12. The triangles ABC, $A'B'C'$ have the angles at C and C' equal to right angles, all the other angles being acute; further, the sides AB, $A'B'$ are equal. But the angle CAB is greater than the angle $C'A'B'$. To prove that CB is greater than $C'B'$, AC is less than $A'C'$, and the angle ABC is less than the angle $A'B'C'$.

Through A draw the straight line AX so that the angle BAX is equal to the angle $B'A'C'$, then AX cuts BC at some point E

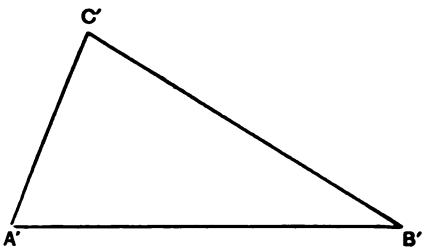


Digitized by Google

between B and C; then by Art. 11 the angle AEC is acute, and therefore the angle BEX is acute; also the angle BAE is acute.

Therefore the perpendicular from B on AX falls on AE produced through E.

Call it BD; then the triangles BAD, $B'A'C'$ are congruent (Art. 6).



Therefore AD is equal to $A'C'$, BD is equal to $B'C'$.

The angles ABD, $A'B'C'$ are equal; hence $A'C'$ is equal to AD, which is greater than AE and this is greater than AC by Art. 11.

$B'C'$ is equal to BD which is less than BE by Art. 11, and therefore less than BC.

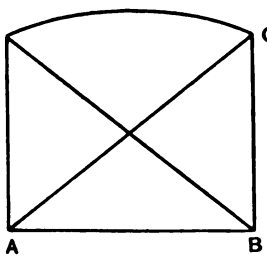
Also the angle $A'B'C'$ is equal to the angle ABD, which is greater than the angle ABC.

Second Group of Theorems.

Some Important Quadrilaterals.

Theorem X.

13. Saccheri's Quadrilateral. Let $ABCD$ be a quadrilateral, let the angles at A and B be right angles, and let AD be equal to BC, to prove that the angles at C and D are equal. (The proof is left to the reader.)



Saccheri now makes three hypotheses. If the angles at C and D of the quadrilateral $ABCD$ are each equal to a right angle he calls this the hypothesis of the right angle.

If they are both obtuse angles he calls this the hypothesis of the obtuse angle; and in like manner if the equal angles are both acute angles he calls this the hypothesis of the acute angle.

Digitized by Google

Theorem XI.

14. Lambert's Quadrilateral. If F, G be the middle points of DC, AB respectively in the figure of Saccheri's quadrilateral (Art. 13), then GF cuts AB, CD at right angles. (The proof is left to the reader.)

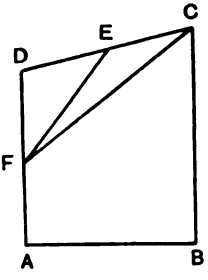
Also $GBCF$, $GADF$ are two quadrilaterals, each having three right angles. These are called Lambert Quadrilaterals. Such a quadrilateral has three right angles and one obtuse angle when the hypothesis of the obtuse angle holds.

Theorem XII.

15. In the quadrilateral $ABCD$ let the angles at A and B be right angles and let the angle ADC be greater than the angle BCD, it is required to prove that AD is less than BC; and conversely if AD is less than BC it is required to prove that the angle ADC is greater than the angle BCD.

(Before proving these propositions it should be observed that they are related to the propositions corresponding to Euc. I. 18 and 19 in Elliptic Geometry. In fact, by producing DA and CB to meet at K we get a triangle KCD, and then applying articles 9 and 10 we find, when ADC is greater than BCD, that KC is greater than KD; from which it follows that BC is greater than AD. I now proceed to give proofs of the above propositions such as Saccheri might have given if he had not made use of Euc. I. 16.)

In the quadrilateral $ABCD$ the angles at A and B are right angles and the angle at D is greater than the angle at C; it is required to prove that AD is less than BC. [It will be supposed that AD, BC are each less than L, and AB is less than $2L$.]



Bisect DC at E and draw through E a perpendicular to DC.

(i) It will be proved in the first place that this perpendicular cannot cross AD between A and D.

If possible let it cross AD at F. Join FC, FE.

The angles FDE, FCE are equal; therefore FDE is less than BCD, i.e. ADC is less than BCD, which is contrary to the hypothesis.

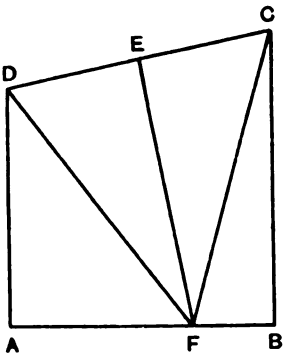
(ii) Suppose next the perpendicular through E to CD crosses AB at F between A and B; then the angles FDE, FCE are equal, but the angle ADC is greater than the angle BCD.

Therefore the angle ADF is greater than the angle BCF; also

Digitized by Google

FD is equal to *FC*: hence, in the right-angled triangles *ADF*, *BCF* the hypotenuses are equal, then if the other angles are acute, it follows by Art. 12 that *AD* is less than *BC*.

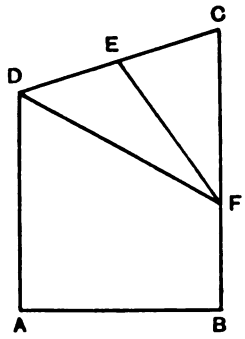
[Saccheri would not have doubted that the other angles of the right-angled triangle *FAD*, *FBC* were acute.



In Elliptic Geometry it follows from the restriction that *AD* and *BC* are each less than *L*, that the angles *AFD*, *BFC* are acute.

Now *FD* and *FC* being equal cannot be greater than *L*; for if they were, then *FA* and *FB* would each be greater than *L*. And therefore the length *AFB* would be greater than 2*L*, contrary to the hypothesis. Hence *FD*, *FC* are each less than *L*.

Then *AD* and *BC* being each less than *L*, it follows that *FA* and *FB* are each less than *L*. Therefore the angles *ADF*, *BCF* are each acute, as has been assumed above.]



(iii) Lastly, suppose that the perpendicular through *E* to *CD* crosses *CB* at *F* between *C* and *B*, then *CF* is equal to *DF*. Therefore *CB* is equal to the sum of *DF* and *FB* which are together greater than *DB* which is greater than *DA*. Therefore *BC* is greater than *AD*.

The proof of the second part of the theorem can be obtained from the first part by a reductio ad absurdum.

Third Group of Theorems.

Properties of two straight lines meeting a third at right angles.

Theorem XIII.

16. Let *ABCD* be a Lambert Quadrilateral.

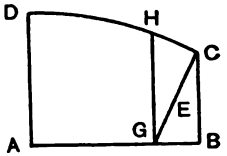
Let the angles at *A*, *B*, *D* be right angles and the angle at *C* obtuse. Suppose that *AD* is less than *L*. Then, by Art. 15, since the angle at *C* is greater than the angle at *D*, *BC* is less than *AD* and therefore less than *L*.

At *C* draw *CE* perpendicular to *DC*.

Then, since *ADC*, *DCE* are right angles, Saccheri would have said that *DA*, *CE* cannot meet.

Digitized by Google

[We must, however, consider the possibility in Elliptic Geometry of CE meeting AD between D and A at some point F. In that case DF would be equal to L. And so AD would be greater than L, contrary to the hypothesis.]



Now, since the angle BCD is obtuse and CE is at right angles to CD, it must cross the boundary DA, AB somewhere. It has been shown that it does not cross DA between D, A. Hence it must cross AB at some point G between A and B.

Since BC is less than L, therefore in Elliptic Geometry the angle BGC is acute and therefore less than the angle GBC, and therefore GC is greater than CB. Draw GH perpendicular to AB to meet DC at H. Since the angles BGC and GCB are both acute, therefore GB, BC are both less than L. Therefore it is known in Elliptic Geometry that GC is less than L. Hence the angle GHC is acute; also the angle HGC is acute, therefore HC is less than L.

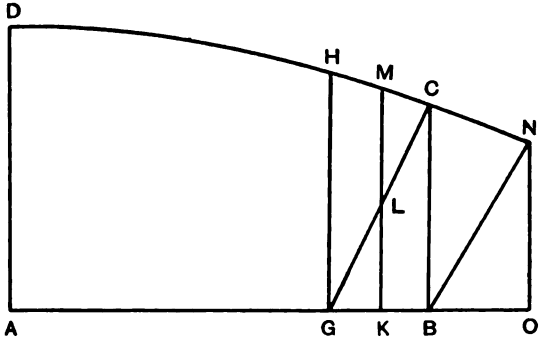
Since HC, GC are both less than L, and GCH is a right angle; therefore GH is less than L.

Since GHC is an acute angle, GH is greater than GC, which is greater than BC. The angle GHD is obtuse and therefore greater than the angle ADH.

Therefore AD is greater than GH (Art. 15) which is greater than BC.

It is now possible, starting from H, to proceed as we did when we started from C.

The process could only come to an end if the straight lines corresponding at any stage to CG, GH coincided, and that could only happen at D.



Now draw KLM perpendicular to GB at K meeting GC at L and HC at M; LC is less than GC, which is less than L. Therefore LMC is acute; then KM is equal to the sum of KL and LM, which

Digitized by Google

is greater than the sum of KL, LC; which is greater than KC, which is greater than BC.

Also KM is less than GM which is less than GH since GCH is a right angle (Art. 11).

Therefore GH is greater than KM which is greater than BC.

Hence in a Lambert Quadrilateral the perpendiculars drawn to the base AB steadily increase as the foot of the perpendicular passes from B to A.

Now suppose DC produced through C; through B draw BN perpendicular to DC meeting it at N; draw NO perpendicular to AB produced through B meeting it at O.

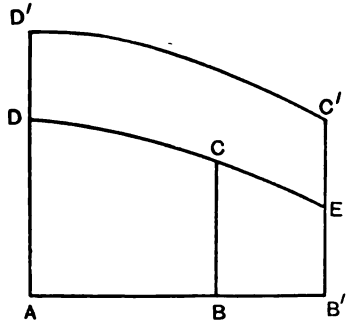
Then the angle OND is obtuse and $AOND$ is a Lambert Quadrilateral with the angle at N obtuse; hence we see that if two straight lines AB, DC meet the straight line AD at right angles, then perpendiculars drawn to one of them, say, AB cut the other DC so as always to make a Lambert Quadrilateral with an obtuse angle; and further that the side perpendicular to AB having one end at the obtuse angle steadily diminishes as the foot of the perpendicular recedes from A.

Theorem XIV.

17. If there be in the plane one Lambert Quadrilateral with an obtuse angle, then all other quadrilaterals in the plane having three right angles must have the remaining angle obtuse.

In the quadrilateral $ABCD$ let the angles at A, B, D be right and the angle at C obtuse; on AB set off AB' equal to a side of the other quadrilateral, and along AD set off a length AD' equal to another side of the second quadrilateral and so complete a figure congruent with the second quadrilateral.

Let DC meet $B'C'$ at E; then (Art. 16) the angle CEB' is obtuse and also by the same article the angle $D'C'B'$ is obtuse, so the second quadrilateral, to which $AB'C'D'$ is congruent, has three right angles and one obtuse angle.



Digitized by Google

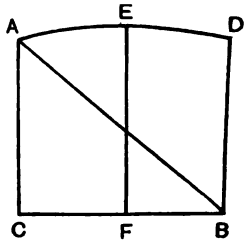
Theorem XV.

18. If the hypothesis of the obtuse angle hold, then the sum of the angles of any triangle is greater than two right angles.

Since any triangle can be divided up into the sum of two right-angled triangles, it is sufficient to prove the theorem for right-angled triangles only. If the right-angled triangle has one of its other angles equal to or greater than a right angle, the proposition is evident.

It is sufficient, therefore, to consider the case of a right-angled triangle in which the other two angles are both acute.

In the triangle ABC let C be a right angle and the angles at A and B both acute; bisect BC at F, draw BD perpendicular to BC and equal to CA; join AD, then $ACBD$ is a Saccheri Quadrilateral.

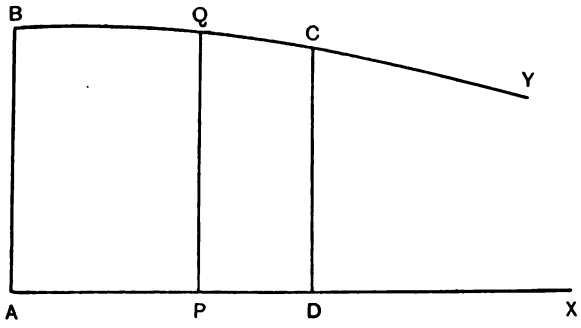


Bisect AD at E and join EF; then $EFBD$ is a Lambert Quadrilateral and the angles BDA, CAD are obtuse. In the figure $EACF$ the angles at E and F are right angles, the angle at A which is obtuse is greater than the angle at C which is a right angle: therefore by Art. 15 AE is less than FC and therefore AD is less than BC.

Hence in the triangles ABC, ABD the sides AB, AC are respectively equal to the sides AB, BD and the side BC is greater than the side AD; therefore the angle CAB is greater than the angle ABD (Euc. I. 25), and therefore the angles of the triangle ABC are together greater than the angles ABD, ABC and ACB, that is, they are greater than two right angles.

Theorem XVI.

19. If the straight lines AX, BY meet the straight line AB at right angles and if, through any point P on AX, a perpendicular be drawn to AX to meet BY at Q, to prove that as P moves along AX away from A the obtuse angle PQB will increase.



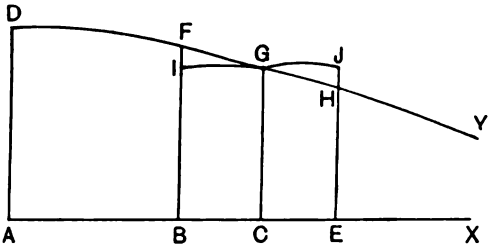
Digitized by Google

Let D be on AX farther from A than P, and let DC perpendicular to AX meet BY at C, then the quadrilateral $PDCQ$ can be divided into two triangles and therefore the sum of its angles is greater than four right angles; but the angles at P and D are right angles, therefore the angles PQC, DCQ are together greater than two right angles: but PQC and PQB are together equal to two right angles.

Therefore the angle DCQ is greater than the angle PQB.

Theorem XVII.

20. Suppose that the straight lines AX, DY cut the straight line AD at right angles. To prove that, if the hypothesis of the obtuse angle hold, they must meet.



On AX take points B, C, E so that BC and CE are equal and each is less than L, and that the perpendiculars at B, C, E to AX meet DY at F, G, H.

It has been proved that AD, BF, CG, EH are in descending order of magnitude, and that the angles EHG, CGF, BFD are also in descending order of magnitude, all of them being obtuse. On BF take BI equal to CG, and on EH produced take EJ equal to CG.

The quadrilaterals $BCGI$, $CEJG$ are congruent.

The obtuse angles BIG, CGI, CGJ and EJG are equal.

The sum of the angles FGI, IGC, CGH is equal to two right angles.

The sum of the angles JGH, CGH is equal to JGC which is equal to IGC.

Hence the excess of JGH over FGI is equal to twice the excess of IGC over a right angle: therefore, since IGC is obtuse it follows that the angle JGH is greater than the angle FGI.

Now GJ is equal to GI, and the angle JGH is greater than the angle FGI, and the angle GJH, which is obtuse, is greater than the angle FIG, which is acute.

Hence if JN be drawn making the angle GJN equal to the angle FIG, and if GN be drawn making the angle JGN equal to

Digitized by Google

the angle *FGI*, the point *N* will be inside the triangle *GJH* and the triangles *GNJ*, *GFI* are congruent. So that *JN* is equal to *FI*.

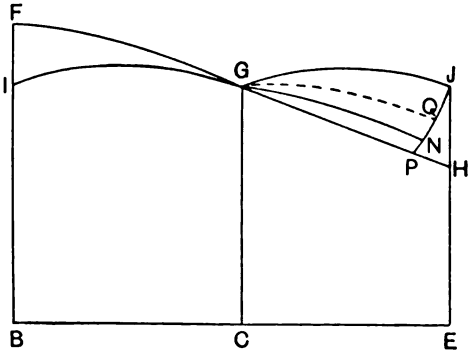
Also *GJN* is equal to *FIG*, which is acute.

And *GNJ* is equal to *GFI*, which is acute.

Hence, if a perpendicular *GQ* be drawn from *G* on *JN*, its foot *Q* falls between *J* and *N*.

Now *GQ* is less than *GJ* since the angle *GJN* is acute.

But in the Saccheri Quadrilateral *GJ* is less than *CE*, since *GJE* is obtuse, and therefore *GJ* is less than *L*, for *CE* by hypothesis is less than *L*.



Therefore *GQ* is less than *L*.

Therefore in the right-angled triangle *GQP* the angle *GPQ* is acute. Therefore the angle *JPH* is obtuse, but *JHP* is acute. Therefore *JH* is greater than *JP*, which is greater than *JN*. Now *JN* is equal to *FI*; therefore *JH* is greater than *FI*.

Therefore *EJ* less *EH* is greater than *BF* less *BI*.

Therefore *CG* less *EH* is greater than *BF* less *CG*.

Thus as we go along *AX* by equal lengths, the decrease in the length of the perpendiculars from points of *DY* on to *AX* increases at a greater rate.

However small the difference in the lengths of the first and second perpendiculars may be, we must be able, if we continue to take equal lengths on *AX*, to erect perpendiculars to *AX* such that the sum of the decreases in the lengths of consecutive perpendiculars will become greater than *AD*. This is impossible if the lines *AX*, *DY* do not meet.

Hence the lines *AX*, *DY* must meet; and as Lambert points out (p. 192) on both sides of *A*.

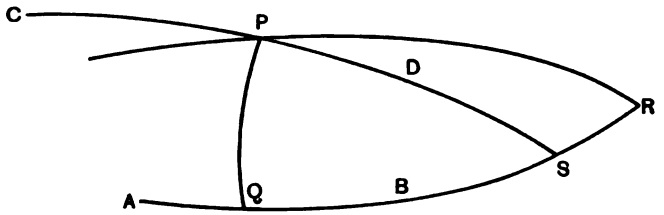
So that two straight lines would enclose a space. This Lambert rejects as absurd. It represents, however, the consequence of the hypothesis of the obtuse angle.

Digitized by Google

Theorem XVIII.

21. It has been proved that two straight lines *which cut a third line at right angles* always meet in two points. Suppose now that AB and CD are *any* two straight lines in the plane; to prove that they also meet in two points.

From any point P on CD draw PQ perpendicular to AB.



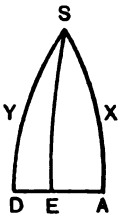
(i) If the angle QPD is a right angle, then by Art. 20 the straight lines CD and AB meet in two points.

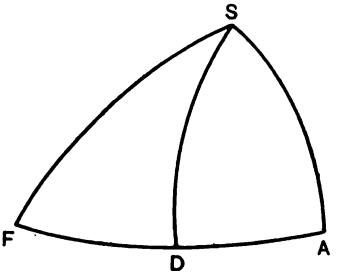
(ii) If the angle QPD is not a right angle, then either it or the angle QPC is acute. Suppose that QPD is acute; at P draw a perpendicular to QP. It was shown in Art. 20 that this perpendicular would meet QB in two points. Call that one of these two points, which is on the same side of PQ as D, is R. Then PD meets QBR in some point S; so the straight lines CD, AB meet in one point S at least.

In order to prove that CD and AB meet in a second point, it is necessary to go somewhat beyond the ideas of Lambert.

We come in fact to the modern treatment of Elliptic Geometry.

We have shown that if AX, DY are both perpendicular to AD then they meet in two points. Call one of these S and join S to any point E on AD between A and D; then if SE is not perpendicular to AD suppose that the angle SEA is acute and therefore the angle SED is obtuse; then it will follow from the triangle SEA which has a right angle at A and an acute angle at E that SE is greater than SA, and from the triangle SED that SE is less than SD; but this is impossible because SA is equal to SD; thus any





Digitized by Google

straight line drawn through S between SA and SD meets AD at right angles. We may by producing AD through D take a point F so that DF is equal to DA.

If we now join SF, the triangles SDA, SDF are congruent; so any straight line through S between SD and SF will meet DF at right angles. Continuing in this way we can show that any straight line whatever through S will cut AD or AD produced at right angles; and consequently it is possible to draw through S two lines making any given angle with each other which are cut at right angles by a third straight line. Now we proved that the two straight lines AB, CD met in one point and we have now proved that at whatever angle they meet they can be cut by a third line at right angles, and therefore by Art. 20 they must meet in two points.

Thus the conclusion is reached that the Geometry of the Plane when the Hypothesis of the Obtuse Angle holds is identical with Elliptic Geometry of the spherical kind.

Digitized by Google

On the substitution of Wallis's postulate of similarity for Euclid's postulate of parallels. Addendum. By Professor M. J. M. Hill, Peterhouse.

[*Received* 25 November, *read* 7 December, 1925.]

The following is an alternative to Proposition II of the former paper*. It does not assume the second of the initial assumptions in Art. 2, nor does it assume that the angles supplementary to equal angles are equal, which Hilbert regards as a proposition (*Foundations of Geometry*, p. 18 of the English translation).

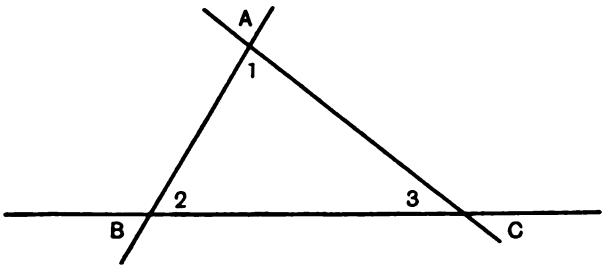
The assumptions in the following proof are:

(i) Through any two points one and only one straight line can be drawn.

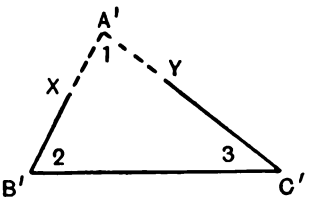
(ii) At any point a straight line can be drawn making with a given straight line through that point on a given side an angle equal to a given angle.

(iii) In the triangle ABC a straight line through A in the angle BAC must meet BC between B and C.

(iv) and (v) are included in Wallis's Postulate of Similarity and may be stated thus:



Let there be three straight lines in a plane which intersect so as to form a triangle ABC. Take any two points B', C'. At B' make the angle $C'B'X$ equal to the angle CBA. At C' make the angle $B'C'Y$ in the same plane as $C'B'X$ equal to the angle BCA.



Then it is assumed

(iv) that $B'X$ and $C'Y$ will meet at some point A', and

(v) that the angle $B'A'C'$ will be equal to the angle BAC.

* *Proc. Camb. Phil. Soc.* vol. 22 (1925), pp. 964–9.

Digitized by Google

It will now be shown that if a be a straight line in a plane, and A a point in the plane, but not on a, then it is possible to draw only one straight line through A in the plane, which does not meet a.

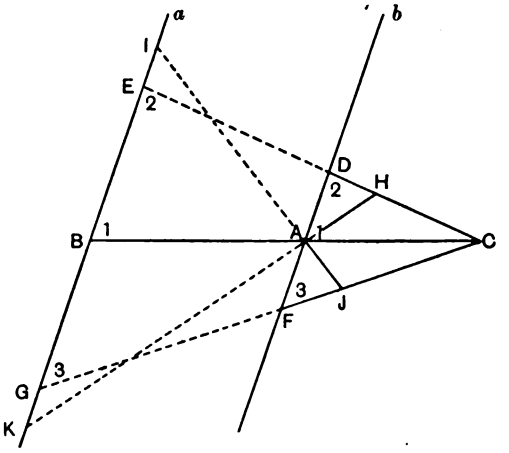
Join A to any point B on a. Produce BA through A to any point C.

Then through A a straight line b can be drawn making with BAC an angle equal to the angle which a makes with BAC.

By Proposition I, b does not meet a.

It will now be shown that every straight line through A, other than b, must meet a.

On b take any point D on the same side of BAC as that on which the equal angles bAC, aBC are situated; and join C to D.



By Art. 4, Corollary (ii), since the angles CAD, CBa are equal, it follows that CD, if produced, must meet a in some point E, and the angles CDA, CEB will be equal, since the triangles CAD, CBE are, by assumptions (iv) and (v) above, equiangular.

On b take any point F on the opposite side of A to D, and join CF.

Then since the angles CDF, CEB are equal, it follows by Art. 4, Corollary (ii) that CF, if produced, will meet a in some point G, and the angles CFD, CGE will be equal, since the triangles CFD, CGE are, by assumptions (iv) and (v) above, equiangular.

Now any straight line through A, other than b, must pass through either the angle CAD or the angle CAF. If it pass through the angle CAD it must meet CD in some point H between C and D.

Then since the angles HDA, HEB are equal, it follows by Art. 4, Corollary (ii) that HA must, if produced, meet a in some point K.

Digitized by Google

If, however, a straight line through A pass through the angle CAF it must meet CF in some point J between C and F.

Then since the angles JFA, JGB are equal, it follows by Art. 4, Corollary (ii) that JA must, if produced, meet a in some point I.

Consequently every straight line through A, other than b, meets a.

NOTE. The triangles CFA, CGB have two angles of the one respectively equal to two angles of the other.

Therefore by Art. 4, Corollary (i) the remaining angles CAF, CBG are equal. These are the angles supplementary to the equal angles CAD, CBE.

This result is proved by Hilbert (*Foundations of Geometry*, p. 18) by a proof in which the congruence of triangles is used.

Digitized by Google

The reciprocation of one quadric into another. By Professor H. F. BAKER.

[*Received* 16 October, *read* 26 October, 1925.]

It is a familiar fact that if two quadric forms, in $(n+1)$ homogeneous variables, be each expressible as a sum of squares, of $(n+1)$ independent linear functions of the variables, then they are polar reciprocals of one another, in regard to any one of 2^n quadrics. The question arises whether this is true for any two quadrics. Segre* states that this is an unsettled question. A solution is given, however, by Terracini† for any two non-degenerate quadrics, supposed to have been reduced to the Weierstrass canonical form. The present note has the purpose of pointing out that a solution is derivable from a remark made by Frobenius‡; this requires a knowledge of the roots of the equation satisfied by the matrix of the two quadrics§.

§ 1. We consider two quadrics in $(n+1)$ homogeneous variables, both of non-vanishing discriminant. The matrix of the coefficients of one of these being m^{-1}, the matrix of the coefficients of the tangential-form of this will be m; let the matrix of the coefficients of the point-form of the other quadric be M. If the quadric which is sought, by means of which the two given quadrics are polar reciprocals of one another, be of matrix p, we are to substitute, for the tangential coordinates u in the first form, mu^2, the values given by $u = px$, and the result is to be the second form Mx^2. Now

$$mu^2 = mpx.px = \bar{p}mp.x^2,$$

where $\bar{p}$ denotes, as usual, the matrix obtained from p by transposition of rows and columns. In other words, the problem is, given m and M, to find a *symmetrical* matrix p such that

$$pmp = M;$$

by hypothesis m and M are symmetrical.

* *Encykl. Math.* III, C. 7, p. 864, 1912.

† *Ann. d. Mat.* XXX, 1921, p. 155.

‡ *Berlin. Sitzungsber.* 1896, p. 7. Frobenius refers to Kelland and Tait, *Quaternions* (1873), chap. X, and to Sylvester, *Papers*, III, pp. 562–567 (1882). Sylvester refers to Babbage's *Calculus of Functions*.

§ This equation may be the same for cases of different invariant factors. For example, if m be the matrix of the coefficients of the form $2xy + z^2 + t^2$, and M_1, M_2 be the respective matrices of the coefficients of the two forms

$$2\theta xy + x^2 + \theta z^2 + \phi t^2, \quad 2\theta xy + x^2 + \phi z^2 + \phi t^2,$$

both the matrices $M_1 m$, $M_2 m$ satisfy the equation

$$(N-\theta)^2 (N-\phi) = 0.$$

Digitized by Google

§ 2. Now we can, without loss of generality, suppose the quadric form $m^{-1}x^2$ reduced to a sum of $(n+1)$ independent squares, at starting, and find the form to which Mx^2 is reduced by the same transformation. Thus it is clear that there is no loss of generality in supposing $m = 1$. Then the problem is to find p such that

$$p^2 = M.$$

But, if the equation satisfied by M be of order $(r+1)$, it is an obvious suggestion of the theory of algebraic numbers to attempt to solve this equation by substituting for p a polynomial in M of order r; the number of conditions to be satisfied by the coefficients in this polynomial is then equal to the number of these coefficients. Frobenius (*loc. cit.*) has given an explicit formula for this polynomial when the roots of the equation satisfied by M are known; this may be regarded as the well-known Lagrange's interpolation formula extended to the case of repeated roots. It is clear that if M is symmetrical, this polynomial will likewise represent a symmetrical matrix. The formula is quoted below (§ 5).

§ 3. Without this reduction of the matrix m to a unit matrix, the equation $pmp = M$ may be solved by finding q such that

$$q^2 = Mm,$$

and then taking $p = qm^{-1}$. To prove this it is only necessary to show that qm^{-1} is a symmetrical matrix; for, from $q^2 = Mm$, we have

$$pmpm = Mm,$$

and hence

$$pmp = M.$$

Now the equation $q^2 = Mm$ leads, because m, M are symmetrical, to

$$\bar{q}^2 = mM;$$

thus, if

$$q = A_r(Mm)^r + \ldots + A_1(Mm) + A_0,$$

we have, for $\bar{q}$,

$$\bar{q} = A_r(mM)^r + \ldots + A_1(mM) + A_0;$$

and hence, as

$$m(Mm)^k = mMmMm \ldots Mm = (mM)^k m,$$

we have

$$mq = \bar{q}m,$$

or

$$qm^{-1} = m^{-1}\bar{q},$$

which, since m is symmetrical, shows that qm^{-1} is symmetrical.

We may illustrate this procedure by deducing Terracini's result. The forms to which Weierstrass reduces any two quadrics have matrices such as

$$\begin{vmatrix} a_1 & . & . & . \\ . & a_2 & . & . \\ . & . & . & . \\ . & . & . & . \end{vmatrix}, \quad \begin{vmatrix} b_1 & . & . & . \\ . & b_2 & . & . \\ . & . & . & . \\ . & . & . & . \end{vmatrix},$$

Digitized by Google

wherein a_1 is a matrix whose diagonal coincides with that of the whole matrix, and b_1 a similarly placed matrix of the same order; likewise a_2, b_2 are similarly placed matrices of equal order; and so on; each of the couples (a_1, b_1), (a_2, b_2), ... is associated with one of the elementary divisors of the matrix formed by the original family of quadrics, that is, of the matrix $M + \rho m^{-1}$. As the product of two matrices of these reduced forms is the matrix

$$\begin{vmatrix} a_1b_1 & . & . & . \\ . & a_2b_2 & . & . \\ . & . & . & . \\ . & . & . & . \end{vmatrix},$$

it is clear on consideration that it is sufficient to consider only the two matrices a_1, b_1, the other couples being similarly treated. Thus we may suppose, in the equation $pmp = M$, limiting ourselves to four rows and columns, and recalling the forms obtained by Weierstrass, that

$$m = \begin{vmatrix} . & . & . & 1 \\ . & . & 1 & . \\ . & 1 & . & . \\ 1 & . & . & . \end{vmatrix}, \quad M = \begin{vmatrix} . & . & 1 & \theta \\ . & 1 & \theta & . \\ 1 & \theta & . & . \\ \theta & . & . & . \end{vmatrix},$$

the elements not written being zeros. Here $m^{-1} = m$, and we have

$$Mm = \begin{vmatrix} \theta & 1 & . & . \\ . & \theta & 1 & . \\ . & . & \theta & 1 \\ . & . & . & \theta \end{vmatrix}.$$

Thus $(Mm - \theta)^4 = 0$; therefore, denoting $Mm - \theta$ by N, a matrix, q, such that $q^2 = \theta + N$, may be obtained by expanding $(\theta + N)^{\frac{1}{2}}$, in powers of N, as far as the term in q^3. Namely we have

$$q = \theta^{\frac{1}{2}}(1 + \tfrac{1}{2}\theta^{-1}N - \tfrac{1}{8}\theta^{-2}N^2 + \tfrac{1}{16}\theta^{-3}N^3);$$

this, however, is

$$q = \theta^{\frac{1}{2}}\left[\begin{vmatrix} 1 & . & . & . \\ . & 1 & . & . \\ . & . & 1 & . \\ . & . & . & 1 \end{vmatrix} + \tfrac{1}{2}\theta^{-1}\begin{vmatrix} . & 1 & . & . \\ . & . & 1 & . \\ . & . & . & 1 \\ . & . & . & . \end{vmatrix} - \tfrac{1}{8}\theta^{-2}\begin{vmatrix} . & . & 1 & . \\ . & . & . & 1 \\ . & . & . & . \\ . & . & . & . \end{vmatrix} + \tfrac{1}{16}\theta^{-3}\begin{vmatrix} . & . & . & 1 \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \end{vmatrix}\right]$$

$$= \theta^{\frac{1}{2}}\begin{vmatrix} 1 & \frac{1}{2}\theta^{-1} & -\frac{1}{8}\theta^{-2} & \frac{1}{16}\theta^{-3} \\ . & 1 & \frac{1}{2}\theta^{-1} & -\frac{1}{8}\theta^{-2} \\ . & . & 1 & \frac{1}{2}\theta^{-1} \\ . & . & . & 1 \end{vmatrix};$$

Digitized by Google

thus $p, = qm^{-1}$, is given by

$$p = \theta^{\frac{1}{2}} \begin{vmatrix} \frac{1}{16}\theta^{-3} & -\frac{1}{8}\theta^{-2} & \frac{1}{2}\theta^{-1} & 1 \\ -\frac{1}{8}\theta^{-2} & \frac{1}{2}\theta^{-1} & 1 & . \\ \frac{1}{2}\theta^{-1} & 1 & . & . \\ 1 & . & . & . \end{vmatrix}.$$

Terracini obtains this result by trying, as a form for p, the matrix

$$p = \begin{vmatrix} \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 \\ \lambda_2 & \lambda_3 & \lambda_4 & . \\ \lambda_3 & \lambda_4 & . & . \\ \lambda_4 & . & . & . \end{vmatrix};$$

then, with the proper m, we have

$$pmp = \begin{vmatrix} (1234) & (234) & (34) & (4) \\ (234) & (34) & (4) & . \\ (34) & (4) & . & . \\ (4) & . & . & . \end{vmatrix},$$

where
$$\begin{aligned} (1234) &= (\lambda_1, \lambda_2, \lambda_3, \lambda_4)\,(\lambda_4, \lambda_3, \lambda_2, \lambda_1) &&= 2\lambda_1\lambda_4 + 2\lambda_2\lambda_3, \\ (234) &= (\lambda_2, \lambda_3, \lambda_4)\,(\lambda_4, \lambda_3, \lambda_2) &&= 2\lambda_2\lambda_4 + \lambda_3^2, \\ (34) &= (\lambda_3, \lambda_4)\,(\lambda_4, \lambda_3) &&= 2\lambda_3\lambda_4, \\ (4) &= (\lambda_4)\,(\lambda_4) &&= \lambda_4^2; \end{aligned}$$

to identify this pmp with the matrix M we require

$$\lambda_4^2 = \theta,\ 2\lambda_3\lambda_4 = 1,\ 2\lambda_2\lambda_4 + \lambda_3^2 = 0,\ 2\lambda_1\lambda_4 + 2\lambda_2\lambda_3 = 0,$$

and these give, for $\lambda_1, \lambda_2, \lambda_3, \lambda_4$, the values obtained otherwise above.

§ 4. We may compare the procedures under § 2 and § 3. To put the matrix m in unit form, it is necessary to take a matrix α such that

$$m = \alpha\bar{\alpha};$$

the equation to be solved is then

$$p\alpha\bar{\alpha}p = M,$$

or, if $r = \bar{\alpha}p\alpha$, $$r^2 = \bar{\alpha}M\alpha.$$

Conversely, when r is found from this, p can be found. To identify this with the procedure in § 3, it is necessary to show that

$$\bar{\alpha}\,(M\alpha\bar{\alpha})^{\frac{1}{2}} = (\bar{\alpha}M\alpha)^{\frac{1}{2}}\,\bar{\alpha};$$

this can be proved precisely as it was proved that

$$m\,(Mm)^{\frac{1}{2}} = (mM)^{\frac{1}{2}}\,m.$$

Digitized by Google

§ 5. The formula of Frobenius, referred to above, for the matrix p such that $p^2 = M$, is as follows: Let the equation of least order satisfied by M be

$$(M - \theta)^a (M - \phi)^b \ldots = 0,$$

where $\theta, \phi, \ldots$ are all different. If the determinant $| M - \rho |$ divide by $(\rho - \theta)^l$, and all the first minors divide by $(\rho - \theta)^{l_1}$, then $a = l - l_1$; and similarly for $b, \ldots$*; thus, to find the equation, we may divide the determinant $| M - \rho |$ by the highest common factor (as regards ρ) of the first minors of this determinant, and then replace ρ by M. The process is rational when M is given; the determination of $\theta, \phi, \ldots$ is a subsequent step. Now expand the fraction

$$\frac{M^{\frac{1}{2}}}{(M - \theta)^a (M - \phi)^b \ldots},$$

where the denominator is obtained by omission of the factor $(M - \theta)^a$, as if M were a single number, in powers of $M - \theta$; denote the terms of this expansion, up to the term in $(M - \theta)^{a-1}$ inclusive, with an arbitrary $\pm$ sign prefixed, by $[\theta]$. Then the polynomial in M, consisting of the sum of such terms as

$$[\theta] (M - \phi)^b \ldots,$$

computed for all the roots $\theta, \phi, \ldots$ in turn, is such that, when squared and then reduced by means of the equation satisfied by M, it reduces to M. This polynomial, U, is of order less than the order of the equation satisfied by M. The verification is by expanding $M - U^2$ in powers of $M - \theta$, $M - \phi$, ... in turn, which shows that this divides by the left side of the equation satisfied by M. If k be the number of different roots of $| M - \rho | = 0$, there are 2^k possibilities for U. The product of any two such terms of U as $[\theta] (M - \phi)^b \ldots$ contains all the factors of $(M - \theta)^a (M - \phi)^b \ldots$; thus U^2 is the sum of the squares of the separate terms.

As a simple illustration, suppose

$$M = \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix};$$

then $M^3 = 1$, and we find

$$M^{\frac{1}{2}} = \tfrac{1}{3} (aM^2 + bM + c),$$

where, with $\epsilon_1^2 = \epsilon_2^2 = \epsilon_3^2 = 1$, $\omega^3 = 1$,

$$a = \epsilon_1 + \epsilon_2 + \epsilon_3, \quad b = \epsilon_1 + \epsilon_2\omega + \epsilon_3\omega^2, \quad c = \epsilon_1 + \epsilon_2\omega^2 + \epsilon_3\omega.$$

§ 6. Certainly for some particular values of m and M, the preceding does not give the general solution of the equation $pmp = M$.

* *Proc. Lond. Math. Soc.* xxx, 1899, p. 196.

Digitized by Google

For instance, when $m = M = 1$, and the matrices are, respectively, of orders 2 and 3, we have solutions, with arbitrary parameters,

$$p = \begin{vmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{vmatrix},$$

$$p = (a^2+b^2+c^2)^{-1} \begin{vmatrix} b^2+c^2-a^2, & -2ab, & -2ac \\ -2ab, & c^2+a^2-b^2, & -2bc \\ -2ac, & -2bc, & a^2+b^2-c^2 \end{vmatrix},$$

these being symmetrical orthogonal transformations.

With somewhat more generality, if the matrix, x, of an orthogonal transformation, for which $x\bar{x} = 1$, be capable of the form

$$x = \beta^{-1} p \alpha,$$

where p involves arbitrary parameters, but α and β are definite, the matrix p being symmetrical, the equation $x\bar{x} = 1$ is the same as

$$\beta^{-1} p \alpha \alpha p \bar{\beta}^{-1} = 1,$$

or

$$pmp = M,$$

whereby the symmetric matrix $m, = \alpha\bar{\alpha}$, is changed, with the symmetric matrix p, involving parameters, into the symmetric matrix $M, = \beta\bar{\beta}$. A case of this, the only one existing for matrices of order 3, is

$$x = \tfrac{1}{4} \begin{vmatrix} 1 & 0 & 2 \\ 0 & 2\sqrt{2} & 0 \\ -3i & 0 & 2i \end{vmatrix} \begin{vmatrix} 0 & 0 & 1 \\ 0 & -1 & -m \\ 1 & -m & \tfrac{1}{2}(1-m^2) \end{vmatrix} \begin{vmatrix} 1 & 0 & -i \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & i \end{vmatrix}.$$

This arises from the symmetric orthogonal form quoted above by taking $a = m$, $b = 2\sqrt{2}$, $c = im$. It corresponds to the fact that the conics

$$2xz + y^2 = 0, \quad 2xz + y^2 + z^2 = 0$$

are polar reciprocals not only with regard to

$$2xz + y^2 + \tfrac{1}{2}z^2 = 0,$$

but also with regard to

$$2xz + y^2 + \tfrac{1}{2}z^2 - \tfrac{1}{2}(mz + 2y)^2 = 0,$$

for any value of m.

§ 7. We have deduced Terracini's result from Frobenius' work. But the converse view may be taken. And, by a combination of both, we are in a position to write down at once a general mth root of any matrix, supposed first put in reduced form—namely, by means of components such as that called q in § 3 above, the fractional coefficients being replaced by the coefficients in the expansion of $(1+t)^{\frac{1}{m}}$, and $\theta^{\frac{1}{2}}$ by $\theta^{\frac{1}{m}}$. Or, indeed, to write down explicitly any algebraic function of the matrix.

Digitized by Google

Note on the pedal locus. By Mr J. P. GABBATT, Peterhouse.

[*Received* 6 October, *read* 26 October, 1925.]

While the paper "On the pedal locus in non-euclidean hyperspace"* was in the press, Professor H. F. Baker kindly directed the writer's attention to a reference† from which it appeared that the euclidean case had first been studied by Beltrami‡. After publication, it was discovered that the main subject of Beltrami's paper was the non-euclidean case§. He proves, by analytical methods, theorems which may be stated as follows: Let $A_0, A_1, \ldots, A_n$ denote the vertices of a simplex, $[A]$, in non-euclidean space of n dimensions, and P a point such that the orthogonal projections of P on the walls of $[A]$ lie in a flat, p; then the locus of P is an $(n-1)$-fold, W, of order $n+1$, which is anallagmatic for the isogonal transformation q. $[A]$; the isogonal conjugate of P is the absolute pole of p, and the envelope, w, of p is therefore the absolute reciprocal of W. Beltrami does not note the theorem, fundamental for the geometrical treatment of the subject, that W is the Jacobian of a certain group of $n+1$ point-hyperspheres. He goes on to show that in the euclidean case the locus W breaks up into an n-ic $(n-1)$-fold and the flat at infinity, while the envelope w does not, in general, break up. Finally, he notes that in two dimensions there is also a special non-euclidean case, in which W breaks up into an order-conic and a line, and w into a class-conic and a point; and that the appropriate condition is fundamentally that which is necessary for the degeneration of a certain class-conic into a pair of points. "But what are the points? And what is the corresponding condition satisfied by the absolute conic? This is a question which it would be interesting to resolve." It does not appear that the subject has been pursued further; and in the present note an attempt is made to discuss fully the analogous case of degeneracy in n dimensions.

Let $[A]$ denote a simplex in non-euclidean space of n dimensions; let A_i denote a vertex of $[A]$, and a_i that wall of $[A]$ which does not contain A_i; let Ω denote the Absolute, B_i, b_i the absolute pole, polar of a_i, A_i respectively, and $[B]$ the simplex determined by the points B_i. Now (for each of the n possible values of j) the point B_j and the join of a_i, a_j determine a flat. If B_i' denote the meet of the n flats so determined, then (2·22) B_i' is the isogonal conjugate, q. $[A]$, of B_i. Let $[B']$ denote the simplex determined

* Gabbatt, *Proc. Camb. Phil. Soc.* XXI (1923), 763–771. References in the text, thus: (2·22), are to this paper.

† *Encyk. d. math. Wiss.* III, C 7, 806–7, footnote 121.

‡ *Mem. Acc. Bologna* (3) VII (1876), 241–262 = *Op. mat.* III, 53–72.

§ See *Encyk. d. math. Wiss.* III, C 7, 962.

Digitized by Google

by the points B_i', and b_i' that wall of $[B']$ which does not contain B_i'; then (2·31) there is a quadric, Ω', the *Secondary* q. $[A]$, such that $[A]$, $[B']$ are polar q. Ω' $(i = 0, 1, \ldots, n)$.

Let the locus, W, of a point, P, which is such that the orthogonal projections of P on the walls of $[A]$ lie in a flat, p, be termed the *pedal locus*, and the envelope of p the *pedal envelope*, q. $[A]$*. Then (3·11, 3·22) W is the Jacobian of the point-hyperspheres determined by the $n + 1$ points B_i, and (3·1) contains those points and the points B_i'. The locus W also contains all the $(n - 2)$-folds, $a_i a_j$, of $[A]$, and the $n + 1$ joins, $a_i b_i$, of corresponding walls of $[A]$, $[B]$.

If the $n + 1$ lines $A_i B_i$ meet at a point, C, so that $[A]$ is orthocentric, and C the orthocentre of $[A]$; then the $n + 1$ linear $(n - 2)$-folds $a_i b_i$ lie in a flat, c, the *orthaxial* of $[A]$; the $n + 1$ lines $A_i B_i'$ meet (2·4) at a point, C', the isogonal conjugate, q. $[A]$, of C; and the flat b_i' contains $a_i b_i$ $(i = 0, 1, \ldots, n)$. The locus W is now (5·13) the Hessian of a cubic $(n - 1)$-fold, U†; isogonally conjugate points of W are conjugate poles q. U; and (1·2, 3·12, 5) c is the mixed polar flat, q. U, of any other pair of isogonally conjugate points. We shall need three further theorems on this case (the *orthocentric case*).

First, it is clear from the specification of B_0' that the polar flat of C q. the n-ple of flats (a_0^{n-1}, b_0) is the flat, d_0, determined by B_0' and $a_0 b_0$. Now the isogonal transform, c_0' (q. $[A]$), of the flat b_0' is an n-ic $(n - 1)$-fold containing each of the points $A_1, A_2, \ldots, A_n$ $(n - 1)$-ply and each of the points $B_1, B_2, \ldots, B_n$ simply; thus d_0 is the polar flat of C q. c_0'. Hence, and similarly: *If the simplex* $[A]$ *be orthocentric, and if* C *denote the orthocentre of* $[A]$, *and* c_i' *the isogonal transform, q.* $[A]$, *of the flat* b_i'; *then the polar flat, q.* c_i', *of* C *is the flat determined by the point* B_i' *and the join of the flats* a_i, b_i $(i = 0, 1, \ldots, n)$. (I)

Again, the polar, q. Ω', of the line $A_i B_i'$ is the $(n - 2)$-fold $a_i b_i'$, which in the present (orthocentric) case is $a_i b_i$. Moreover, if S denote the polar quadric, q. U, of C, then (5·12) the polar flat, q. S, of A_i contains $a_i b_i$; and since C, C' are isogonal conjugates, therefore the polar flat, q. S, of C' is the orthaxial, c, which also contains $a_i b_i$. Thus $A_i B_i'$, $a_i b_i$ are polar q. both the quadrics Ω', S $(i = 0, 1, \ldots, n)$; whence: *In the orthocentric case, if* S *denote the polar quadric of the orthocentre q. the cubic* $(n - 1)$*-fold of which the pedal locus is the Hessian; then* S *and the Secondary,* Ω', *touch at all their common points, and those points are in the orthaxial.* (II)

* Beltrami terms W, w respectively the *hypersteinerian* locus and envelope q. $[A]$. This term is used in a different sense by Brambilla, *Rend. Acc. Napoli*, XXXVIII (1899), 144–5.

† Bauer, *Sitz. Akad. München*, XVIII (1888), 337–354, proves this theorem for three dimensions; see Jessop, *Quartic Surfaces*, Cambridge (1916), 189–190.

Digitized by Google

Lastly, let Σ' denote the common section, by the orthaxial, c, of the quadrics S, Ω'. Then since S, which is the polar quadric, q. U, of C, contains every point, M, of Σ'; therefore the polar flat, m, of M contains C. Moreover, since M is a point of c, therefore m touches the polo-n-ic, c' (q. U), of c; also c' is the isogonal transform, q. $[A]$, of c; and the point of contact of m, c' is the isogonal conjugate of M. Thus: *If Σ' denote the section of the Secondary by the orthaxial, and c' the isogonal transform, q. $[A]$, of the orthaxial; then the isogonal transform of Σ' is the locus of the points of contact of the tangent flats to c' from the orthocentre.* (III)

We now proceed to examine the assumption that, in non-euclidean space, a flat, g, can form part of the locus W. Remembering that the section of W by a wall, a_i, of $[A]$ consists of $n+1$ linear $(n-2)$-folds (viz. the joins of a_i with the remaining walls of $[A]$, and with the corresponding wall, b_i, of $[B]$), and having regard to the section of the walls of $[A]$ by the assumed flat, g; we see that g must contain *either* (i) at least two of the joins $a_i a_j$ *or* (ii) at least n of the joins $a_i b_i$. In case (i), g must be a wall, a_i, of $[A]$; this occurs when a_i either touches the Absolute, or is the absolute polar of A_i. The case is not analogous to that of Beltrami, and will not be further discussed.

In case (ii), if the Absolute be non-degenerate, then at least n of the lines $A_i B_i$ meet at a point. The $n+1$ lines $A_i B_i$ therefore meet at a point*; the simplex $[A]$ is orthocentric; and the flat g is the orthaxial, c. Now if c form part of the locus W, then the remainder of the locus is the isogonal transform, c', of c. Also W contains all the points B_i and their isogonal conjugates B_i'. Thus *either* (α) c contains all the points B_i *or* (β) c contains at least one of the points B_i'. In case (α), the Absolute is a flat quadric in c; this is the euclidean case.

We now consider case (β). Since, whenever $[A]$ is orthocentric, the flat determined by B_0' and $a_i b_i$ (which lies in the orthaxial c) contains all the remaining vertices of $[B']$ except B_i'; therefore if c contain (e.g.) B_0', then c also contains all the n points B_i' $(i = 1, 2, \ldots, n)$. The quadric Ω' then degenerates into a flat quadric in c, and is therefore identical with Σ' (see III); also B_i' is the pole, q. Σ', of $a_i b_i$. Thus we have the following: *In non-euclidean space of n dimensions, the pedal locus, q. a simplex $[A]$, includes a flat if $[A]$ is orthocentric, and one of the points B_i' lies in the orthaxial, c, of $[A]$. Then all the $n+1$ points B_i' lie in c; the $n+1$ points B_i lie in the isogonal transform, c', (q. $[A]$) of c; the pedal locus consists of c and c'; and the pedal envelope consists of the orthocentre, C, of $[A]$ and the absolute reciprocal of c'.* (IV)

Referring to the case of theorem IV as *Beltrami's case*, we have from (I): *In Beltrami's case, the polar flat, q. c', of C is c.* (V)

* See Schläfli, *J. f. Math.* LXV (1866), 189–197.

Digitized by Google

Also from (II, III): *In Beltrami's case, the $n+1$ points B_i' are the poles (q. a flat quadric, Σ', in c) of the $n+1$ joins of corresponding walls of $[A]$ and the absolute polar simplex, $[B]$; and Σ' is the isogonal transform, q. $[A]$, of the locus of the points of contact of the tangent flats from C to c'.* (VI)

It has not yet been formally proved that, in the orthocentric case, incidence of one of the points B_i' on the orthaxial is possible. It may therefore be noted that, if the restriction that the quadric Ω is the Absolute be removed, and if the terms required (perpendicular, isogonal conjugate, etc.) be defined projectively q. Ω, then (2·33) the substitution of Ω' for Ω results in the interchange of the simplices $[B]$, $[B']$ and of the quadrics Ω, Ω'; and the isogonal transformation, q. $[A]$, remains unaltered. Moreover (2·42) the simplex $[A]$ is either orthocentric q. both the quadrics Ω, Ω' or q. neither; and (5·14) in the orthocentric case (but only in that case) the pedal locus q. $[A]$ is unaffected by the substitution. Beltrami's case is therefore projectively identical with the orthocentric case in euclidean geometry; the apparent difference (as regards degeneracy) in the behaviour of the pedal envelope in the two cases being due to the circumstance that, in the former case, that one of the two quadrics Ω, Ω' which is non-degenerate is regarded as Absolute, while in the latter, that which is degenerate is so regarded.

It is interesting to remark that the euclidean expression of theorem V is as follows: If c' denote the pedal (n-ic) locus, q. an orthocentric simplex in euclidean space of n dimensions, and C' the isogonal conjugate of the orthocentre; then the polar flat of C' q. c' is the flat at infinity. Now since any line parallel to an edge of the simplex meets c' $(n-2)$-ply at infinity, it follows that: *In euclidean space of n dimensions, if $[A]$ denote any orthocentric simplex, and C' the isogonal conjugate, q. $[A]$, of the orthocentre: then any line containing C' and parallel to an edge of $[A]$ meets the pedal locus (q. $[A]$) at two actual points only; and the segment determined by the two points is bisected at C'**. In a euclidean plane, this reduces to the theorem that the circumcentre of a triangle is the isogonal conjugate of the orthocentre. (VII)

In the actual case discussed by Beltrami, the theorems IV–VI take the following form: *If A_0, A_1, A_2 denote the vertices of a non-euclidean plane triangle, $[A]$, and B_0, B_1, B_2 the corresponding vertices of the absolute polar triangle, $[B]$, and if the six points A_i, B_i be on a conic, c'; then the six sides a_i, b_i of $[A]$, $[B]$ touch a conic*, K, *the absolute reciprocal of c'. The pedal locus, q. $[A]$, breaks up into the conic c' and the orthaxis, c, of $[A]$; the pedal envelope breaks up into the conic* K *and the orthocentre, C, of $[A]$; c' is the isogonal transform, q. $[A]$, of c; and c is the polar, q. c', of C. If B_i' denote*

* Cf. Gabbatt, *Proc. Lond. Math. Soc.* (2), XXIV (1925), 173.

Digitized by Google

the isogonal conjugate, q. $[A]$, *of* B_i, *and* C_i *the meet of* a_i, b_i; *then the two points* B_i', C_i *are conjugates in an involution range, on c, of which the double points, D, D', are the meets of c, c', and constitute the isogonally conjugate point-pair on c* $(i = 0, 1, 2)$*. (VIII)

If the conditions of theorem VIII be satisfied, and if any line containing C meet c' at the points X, Y; then it is easy to show that $A_i(A_jA_kXY) \barwedge B_i(B_jB_kYX)$. In particular, the line XY may be the absolute polar of B_i'; and in that case it may readily be proved that the four lines A_iX, A_iY, B_iX, B_iY all touch the Absolute. The relation $A_i(A_jA_kXY) \barwedge B_i(B_jB_kYX)$ then expresses the congruence of one of the supplementary angles a_ja_k with one of the supplementary angles b_jb_k $(i, j, k = 0, 1, 2)$. Thus: *In a non-euclidean plane, if* $[\bar{A}]$ *denote any triangle of which the angles are, independently, either congruent or supplementary to the angles of the triangle* $[A]$*; then the congruence of one of the triangles* $[\bar{A}]$ *with the absolute polar triangle of* $[A]$ *is a necessary condition for degeneracy (of the type considered) of the pedal locus.* It may be shown† that the congruence, for a *single* pair of values of j, k, of one of the supplementary angles a_ja_k with one of the supplementary angles b_jb_k is in general a sufficient condition for degeneracy. (IX)

In the special case of theorem VIII, the point-pair D, D' replaces the class-conic, Ω', q. which, in general, the triangles $A_0A_1A_2$, $B_0'B_1'B_2'$ are mutually polar. It is to be observed that Ω' is not the conic of which the degeneracy in the same case is noted by Beltrami. If the Absolute be expressed by the line-equation

$$\Omega \equiv (a_{00}, \ldots, \ldots, a_{12}, \ldots, \ldots \mathcal{X}\xi_0, \xi_1, \xi_2)^2 = 0,$$

then

$$\Omega' \equiv (a_{00}^{-1}, \ldots, \ldots, a_{12}^{-1}, \ldots, \ldots \mathcal{X}a_{00}\xi_0, a_{11}\xi_1, a_{22}\xi_2)^2 = 0,$$

whereas Beltrami's class-conic is expressed by the equation

$$(a_{00}^{-1}, \ldots, \ldots, a_{12}^{-1}, \ldots, \ldots \mathcal{X}\xi_0, \xi_1, \xi_2)^2 = 0;$$

the condition for the case of theorem VIII being the vanishing of the determinant

$$\begin{vmatrix} a_{00}^{-1} & a_{01}^{-1} & a_{02}^{-1} \\ a_{10}^{-1} & a_{11}^{-1} & a_{12}^{-1} \\ a_{20}^{-1} & a_{21}^{-1} & a_{22}^{-1} \end{vmatrix} \quad (a_{ij} = a_{ji}).$$

* Cf. Neuberg, *Mathesis*, (3) VIII (1908), 159–160; Juhel-Rénoy, *ibid.* 257–258.

† E.g. from Gabbatt, *Proc. Camb. Phil. Soc.* XXI (1923), 297–362, §§ 9·12, 15·1.

Digitized by Google

On the series of forms of Jacobi's rotating liquid ellipsoid. By W. F. SEDGWICK, M.A., Trinity College.

[*Received* 20 October, *read* 26 October, 1925.]

The object of these notes is twofold. First, to connect the theory of the possible shapes of Jacobi's ellipsoidal equilibrium figure, for a given mass of homogeneous rotating liquid, with the theory of its stability and other properties given by Prof. H. F. Baker in his paper "On the stability of rotating liquid ellipsoids" in *Proc. Camb. Phil. Soc.* vol. 20, pp. 190–97. And secondly, to indicate two inconclusive reasonings on the subject in Thomson and Tait's *Natural Philosophy* (2nd edition), arts. 778 and 778′, and to complete the proofs of the results in question.

1. Taking $\sqrt{a}$, $\sqrt{b}$, $\sqrt{c}$ as the semi-axes of the ellipsoid, in descending order of magnitude, Prof. Baker has provided a concise proof, in the paper* referred to, that as $a - b$ increases numerically, $a + b$ (which is proportional to the moment of inertia about the axis of rotation c), the moment of momentum, and the total energy (kinetic + potential), all constantly increase when the volume or, the density being invariable, the mass of the liquid is given; whilst the angular velocity, on the other hand, constantly decreases. For the fact that $a - b$ constantly increases, as we proceed along the series of ellipsoids, he refers to the numerical tables giving the values of the axes at successive intervals. As, however, it is difficult in such cases, apart from strict mathematical demonstration, to feel quite satisfied that in the course of the calculations some additional series of figures, or some minor maxima or minima resulting from variations extending only over short intervals, may not have been accidentally overlooked, it seems desirable to connect with Prof. Baker's equations a rigorous proof that, as a constantly increases, both b and c constantly decrease. And this can easily be done.

2. Writing

$$\left.\begin{aligned} \phi &= \int_0^\infty \frac{dx}{\{(x+a)(x+b)(x+c)\}^{\frac{1}{2}}}, \quad \frac{1}{h} = a + b, \\ \phi_{12} &= \frac{\partial^2 \phi}{\partial h \partial c}, \quad \phi_{22} = \frac{\partial^2 \phi}{\partial c^2}, \end{aligned}\right\} \ldots\ldots(1)$$

* In this paper, on p. 191, line 12 from foot, for "$(\xi_1^2 - \xi_1\eta_1 + \eta_1^2)$" read "$u\,(\xi_1^2 - \xi_1\eta_1 + \eta_1^2)$"; on p. 192, line 13 from foot, for "$\phi_1 = c$" read "$\phi_1 = \sigma$"; and on p. 194, line 6 from foot, for "$b + x$" read "$b + y$."

Digitized by Google

Prof. Baker shows (p. 194, *l.c.*) that

$$\phi_{12} = \tfrac{1}{2}ch P, \quad \phi_{22} = -h^2 P - R, \qquad \text{.....(2)}$$

where P, R are each essentially positive and different from zero in the problem under consideration, so that ϕ_{12} is positive, and ϕ_{22} negative. Also we have the equation (p. 193, *l.c.*)

$$h^2(\delta a + \delta b) = -\delta h = \frac{\phi_{22}}{\phi_{12}}\delta c, \qquad \text{.....(3)}$$

showing at once that if c constantly diminishes, h does so also, and therefore $a + b$ constantly increases.

Now by the condition of constant mass, i.e. of volume, we have

$$-b\delta a - a\delta b = \frac{ab}{c}\delta c. \qquad \text{.....(4)}$$

From (3) and (4)

$$(a-b)h^2\delta a = \frac{a\delta c}{c\phi_{12}}(c\phi_{22} + bh^2\phi_{12}), \qquad \text{.....(5)}$$

$$-(a-b)h^2\delta b = \frac{b\delta c}{c\phi_{12}}(c\phi_{22} + ah^2\phi_{12}). \qquad \text{.....(6)}$$

But by (2), since $ah + bh = 1$,

$$c\phi_{22} + bh^2\phi_{12} = c\phi_{22} + 2h\phi_{12} - h(1+ah)\phi_{12} = -cR - h(1+ah)\phi_{12},$$

and $c\phi_{22} + bh^2\phi_{12}$, and similarly $c\phi_{22} + ah^2\phi_{12}$, are accordingly negative. Thus δa and δb always have opposite signs; and, when $a > b$, δb and δc have each the opposite sign from δa. Moreover, it is impossible for any one of the three quantities a, b, c to vary without a simultaneous variation in each of the other two; except to the first order of small quantities at the point of junction with Maclaurin's spheroid, where $a = b$ and c is a maximum. Starting from this point of junction, we must suppose either a or b to increase. If a increases, both b and c constantly decrease. The supposition that b increases leads obviously by symmetry to a precisely similar set of figures; for the future, therefore, we shall always suppose that it is a that increases. Since

$$(\sqrt{a} + \sqrt{b})^2 = a + b + 2\sqrt{ab},$$

and by (3) and (4) both $a + b$ and ab constantly increase, it follows that the sum of the axes of the principal section perpendicular to the axis of rotation constantly increases; their difference obviously does so.

3. The Jacobian condition of rotational equilibrium, as derived from the criterion that the total energy is a minimum (p. 192, *l.c.*), may be written, when the mass is constant, in the form

$$\frac{1}{c}\int_0^\infty \frac{\{a(b-c) - bc - cx\}\,x\,dx}{\{(x+a)(x+b)(x+c)\}^{\frac{3}{2}}} = 0. \qquad \text{.....(7)}$$

Digitized by Google

On making use of the identity

$$0 = \frac{2}{c}\int_0^\infty \frac{d}{dx}\left[\frac{x}{\{(x+a)(x+b)(x+c)\}^{\frac{1}{2}}}\right]dx$$

$$= \frac{1}{c}\int_0^\infty \frac{2abc + (bc+ca+ab)x - x^3}{\{(x+a)(x+b)(x+c)\}^{\frac{3}{2}}}\,dx, \qquad \ldots\ldots(8)$$

we can transform (7) into

$$\frac{1}{c}\int_0^\infty \frac{2abc + 2abx - cx^2 - x^3}{\{(x+a)(x+b)(x+c)\}^{\frac{3}{2}}}\,dx = 0,$$

or

$$\frac{1}{c}\int_0^\infty \frac{2ab - x^2}{(x+a)(x+b)\{(x+a)(x+b)(x+c)\}^{\frac{1}{2}}}\,dx = 0. \quad \ldots(9)$$

From (7) we see that as a becomes large, b approaches equality with c, since otherwise the positive part of the integral would become indefinitely large compared with the negative part. Accordingly, since abc is finite, we have $b = c = 0$, when $a = \infty$. The same result follows from (9), taken in conjunction with the fact that c is less than the least of a and b; a fact which follows at once from (7), since the integrand in that equation would be always negative unless $c < \dfrac{ab}{a+b}$.

4. Again, by (2) and (6) we have

$$\frac{1}{n}\delta(b^n - c^n) = c^{n-1}\delta c\left\{-1 - \frac{b^n}{c^n}\cdot\frac{a}{a-b} - \frac{b^n}{c^{n-1}}\cdot\frac{1}{(a-b)h^2}\cdot\frac{\phi_{22}}{\phi_{12}}\right\}$$

$$= c^{n-1}\delta c\left\{-1 - \frac{b^n}{c^n}\cdot\frac{a}{a-b} + \frac{b^n}{c^n}\cdot\frac{2}{(a-b)h} + \frac{b^n}{c^n}\cdot\frac{2}{(a-b)h^2}\cdot\frac{R}{hP}\right\}. \qquad \ldots\ldots(10)$$

Now

$$-\frac{b^n}{c^n}\cdot\frac{a}{a-b} + \frac{b^n}{c^n}\cdot\frac{2}{(a-b)h} = \frac{b^n}{c^n}\cdot\frac{a+2b}{a-b} > \frac{b^n}{c^n} > 1, \quad \ldots(11)$$

provided $n > 0$. Hence, when n is real and positive, $b^n - c^n$ constantly decreases to zero; in particular, both $b - c$ and $\sqrt{b} - \sqrt{c}$ constantly decrease to zero. Since it has been shewn that both $a + b$ and $\sqrt{a} + \sqrt{b}$ constantly increase, it follows that $a + c$ and $\sqrt{a} + \sqrt{c}$ both constantly increase, since

$$a + c = a + b - (b - c), \text{ etc.}$$

Thus it has been proved that in the case of each of the two principal sections through the greatest axis (a), the sum and difference of the axes, the area, and the moment of inertia about the axis perpendicular to the section all constantly increase to infinity, as we proceed along the series of Jacobi's ellipsoids from

Digitized by Google

the junction with Maclaurin's series. On the other hand, all these quantities constantly decrease to zero in the case of the third (smallest) principal section.

5. As regards the ratios of the axes, it is obvious that $\frac{b}{a}$ and $\frac{c}{a}$ constantly diminish to zero as a increases to infinity. Also by (10)

$$\delta\left(\frac{b}{c}\right) = \frac{c\delta(b-c) - (b-c)\delta c}{c^2}$$

$$= \frac{\delta c}{c}\left\{-\frac{b}{c} - \frac{b}{c}\cdot\frac{a}{a-b} + \frac{b}{c}\cdot\frac{2}{(a-b)h} + \frac{b}{c}\cdot\frac{2}{(a-b)h^2}\cdot\frac{R}{hP}\right\},$$

and this is negative by (11). Hence $\frac{b}{c}$ constantly diminishes to unity, when $b = c = 0$. Thus when any one of the three ratios is given, the other two are determinate and unique, since any given value of any one of the ratios occurs only once in the series of ellipsoids. This has been proved on the supposition that the volume, and therefore abc, is constant; but it holds generally. For on writing $x = c\mu$, $a = \frac{c}{s}$, $b = \frac{c}{t}$, in (7), and inserting the factor abc in order that the variational equation may relate to the energy multiplied only by a numerical factor (p. 192, *l.c.*), we find that the equation may be written

$$(abc)^{\frac{1}{2}}\int_0^\infty \frac{(1-s-t-st\mu)\,\mu\,d\mu}{\{(1+\mu)(1+s\mu)(1+t\mu)\}^{\frac{3}{2}}} = 0. \quad \ldots\ldots(12)$$

This is virtually the form of the Jacobian relation used by C. O. Meyer in his original investigation (see § 9 below), by means of which he showed (as corrected by Liouville) that as s increases from 0 to 1, t diminishes from 1 to 0, and *vice versa*; and it is apparent from this form of the relation that it conditions only the form of the ellipsoid, without regard to its volume. Consequently, if for any given value of the ratio of any two of the axes there were more than one value for the other two ratios, the same multiplicity of values would arise also when the volume of the ellipsoid is kept constant. But we have seen that in that case all the ratios are uniquely determined by any one of them; hence this proposition holds generally.

6. We now come to the treatment of the matter in Thomson and Tait's *Natural Philosophy* (§ 778). Taking a, b, c as the semi-axes, and writing $D = (a^2+u)^{\frac{1}{2}}(b^2+u)^{\frac{1}{2}}(c^2+u)^{\frac{1}{2}}$, the authors write the Jacobian relation in the form

$$\int_0^\infty \frac{u\,du}{D^3}\left(\frac{1}{a^2} + \frac{1}{b^2} - \frac{1}{c^2} + \frac{u}{a^2b^2}\right) = 0,$$

Digitized by Google

and remark that it "may be regarded as an equation to determine c^2 for any given values of a and b." They then proceed as follows:—

"It has obviously one and only one real positive root; which is proved by remarking, that while u increases from zero to infinity, u/D^3 decreases continually to zero, and the last factor under the integral sign continuously increases, only reaching a positive value for infinitely great values of u when c is zero, and being positive for all values of u when $1/c^2 = \text{or} < 1/a^2 + 1/b^2$: and that, for any constant value of u, the last factor increases with increase of c^2. As every element of the integral is positive when

$$1/c^2 = \text{or} < 1/a^2 + 1/b^2$$

and as we may write this inequality as follows,

$$c^2 = \text{or} > b^2/(1 + b^2/a^2),$$

we see that if $c = \text{or} <$ the less of a, or b, every element of the integral is positive, and we infer that the root c is less than the least of a or b."

Remark in the first place that the last sign of inequality should be "$>$" instead of "$<$"; and that for "u/D^3" we must read "$1/D^3$," if the passage is to read correctly as it stands, since u/D^3 increases at first. We can in fact easily show that u/D^3 increases to a single maximum: the reasoning would not, however, be affected if this algebraical expression were allowed to stand with suitable alteration of the wording. The argument is that, as c increases from zero, an ever-increasing range of the integral, including the larger values of u, becomes positive, and consequently only an ever-decreasing range, including the smaller values of u, remains negative; whilst the value of the positive elements continually increases, simultaneously with a continuous decrease in the numerical value of the negative elements. Consequently, as the integral is negative at first but finally positive, a stage must be reached, and only one such stage, when the positive and negative parts of the integral balance, which determines a unique value of c^2 or c. Unfortunately the authors appear to have overlooked the fact that the denominator as well as the numerator of the integrand increases with increasing c, on account of the factor $(c^2 + u)^{\frac{3}{2}}$ in the former; and although this factor causes a further decrease in the numerical value of the negative elements of the integral, it is impossible to say without further investigation whether it may not cause so great a concurrent diminution in the value of the positive elements, at certain stages in the continuous increase of c, as to cause the integral to vanish more than a single time.

Digitized by Google

The defect is, however, easily remedied. For if, with the object of including all the parts of the integrand involving c, we write

$$v = \frac{1}{(c^2+u)^{\frac{3}{2}}}\left(\frac{1}{a^2}+\frac{1}{b^2}-\frac{1}{c^2}+\frac{u}{a^2b^2}\right),$$

we find

$$\frac{dv}{d(c^2)} = \frac{1}{(c^2+u)^{\frac{5}{2}}}\left\{\frac{3}{2}\left(\frac{1}{c^2}-\frac{1}{a^2}-\frac{1}{b^2}\right)+\frac{1}{c^2}+u\left(\frac{1}{c^4}-\frac{3}{2a^2b^2}\right)\right\},$$

which is always positive when, as in our problem, u is positive and

$$\frac{1}{c^2} > \frac{1}{a^2}+\frac{1}{b^2}.$$

Thus v always increases with c^2, and the above argument for the uniqueness of c with given a and b becomes applicable. The same result follows of course from § 5 above, since the ratio of c to a follows uniquely when that of a to b is given. It may be remarked that the converse is not true, for with a given c there is conjoined a singly infinite series of pairs of values for a, b: if, however, the volume also is given it follows from § 2 above that c determines a and b.

7. In § 778′ Thomson and Tait proceed to deal with the case of junction with Maclaurin's spheroids, for which $a = b$. They point out that in this case, on writing $f = \sqrt{\frac{a^2-c^2}{c^2}}$, the Jacobian relation reduces to the equation

$$\frac{\tan^{-1}f}{f} = \frac{1+\frac{1}{3}^3 f^2}{1+\frac{1}{3}^4 f^2+f^4}. \qquad \text{.....(13)}$$

The authors then proceed as follows:—

"When f is increased continuously from zero to infinity the left-hand member of this equation diminishes continuously from unity to zero: the right-hand member diminishes also from unity to zero, but diminishes at first less rapidly and afterwards more rapidly than the other. Thus there is one and only one root," etc.

Now on writing w for the left-hand member of the equation, and w' for the right-hand member, we find

$$\frac{dw}{df} = \frac{1}{f(1+f^2)} - \frac{\tan^{-1}f}{f^2}, \qquad \text{.....(14)}$$

$$\frac{dw'}{df} = -6f\frac{1+6f^2+13f^4}{(3+14f^2+3f^4)^2}. \qquad \text{.....(15)}$$

Digitized by Google

When f is small these equations yield

$$-\frac{dw}{df} = \tfrac{2}{3}f - \tfrac{4}{5}f^3 + \dots,$$

$$-\frac{dw'}{df} = \tfrac{2}{3}f - \tfrac{20}{9}f^3 + \dots.$$

And when $f = 1$,

$$-\frac{dw}{df} = -\frac{1}{2} + \frac{\pi}{4} = \cdot 2854\dots,$$

$$-\frac{dw'}{df} = \cdot 3.$$

Thus at first $$-\frac{dw'}{df} < -\frac{dw}{df},$$

and when $f = 1$, $$-\frac{dw'}{df} > -\frac{dw}{df},$$

and the authors' statement is so far confirmed. But if when f is large we write ϕ for $\frac{1}{f}$, where ϕ is small, we have from (14) and (15)

$$-\frac{dw}{df} = -\frac{\phi^3}{1+\phi^2} + \phi^2\left(\frac{\pi}{2} - \tan^{-1}\phi\right)$$

$$= \frac{\pi}{2}\phi^2 - 2\phi^3 + \text{terms in } \phi^5, \text{ etc.},$$

$$-\frac{dw'}{df} = 6\phi^3\frac{13 + 6\phi^2 + \phi^4}{(3 + 14\phi^2 + 3\phi^4)^2}$$

$$= \tfrac{26}{3}\phi^3 + \text{terms in } \phi^5, \text{ etc.}$$

Hence, when f is large we again have $-\frac{dw'}{df} < -\frac{dw}{df}$, as at first; and although when f is large $w > w'$, so that there is certainly one root of (13), we cannot be certain, without further investigation, that the sign of $\frac{dw}{df} - \frac{dw'}{df}$ may not have changed more frequently, and in such wise as to cause the curves $w = 0$ and $w' = 0$ to meet at more than one point.

8. In order to remove this difficulty the simplest method seems to be to isolate the transcendental function in (13), so that it may disappear on differentiation (cf. a letter written by Laplace to D'Alembert in 1777, *Œuvres*, Tom. XIV (1912), p. 349; or *Méc. Cél.* Livre III, ch. iii, Nos. 20 and 21). For this purpose we consider the function

$$y = f\frac{3 + 13f^2}{3 + 14f^2 + 3f^4} - \tan^{-1}f. \qquad \dots\dots(16)$$

Digitized by Google

On differentiating, we find

$$\frac{dy}{df} = \frac{9 + 75f^2 + 155f^4 - 39f^6}{(3 + 14f^2 + 3f^4)^2} - \frac{1}{1+f^2}$$
$$= \frac{16f^4(1 + 3f^2)(1 - f^2)}{(1+f^2)(3 + 14f^2 + 3f^4)^2}.$$

Thus when f is real and positive, as in the present problem, $\frac{dy}{df}$ vanishes only once, viz. when $f = 1$; and, by what has gone before, y is positive when f is small, but negative when f is infinite. Hence equation (16) has one, and only one, positive root, and this root is greater than unity.

The same result follows of course at once from §§ 2 and 5 above, whether the volume of the ellipsoid is or is not given. The series of Jacobi's ellipsoids has therefore only a single point of contact with the series of Maclaurin's spheroids.

9. It is rather remarkable to what a number of incorrect or inconclusive arguments this problem of the series of possible forms of Jacobi's ellipsoid has given rise. In the first place Lagrange had inferred on insufficient grounds that no form with unequal axes is possible at all; this is the more striking in that he had actually obtained the two equations equivalent to (4) of Art. 374 of Lamb's *Hydrodynamics* (5th edition, 1924), from which Jacobi's relation immediately follows (see the *Méc. Anal., Première Partie*, Sect. VII, Art. 26, on p. 192 of Tome I of Bertrand's edition of 1853, where in line 11 for "f" read $\frac{\text{"f"}}{\text{A}}$ twice; see also Bertrand's note IV on p. 405, *l.c.*)*. With ρ and ω for the density and angular velocity, and a, b, c for the semi-axes, Lamb's equations above cited are

$$\left(\alpha_0 - \frac{\omega^2}{2\pi\rho}\right) a^2 = \left(\beta_0 - \frac{\omega^2}{2\pi\rho}\right) b^2 = \gamma_0 c^2, \qquad \text{.....(17)}$$

where $2\pi\rho$ ($\alpha_0 x$, $\beta_0 y$, $\gamma_0 z$) are the component attractions. Lagrange in effect argues that because α_0 is the same function of a, b that β_0 is of b, a it follows from (17) that $a = b$; a position that is untenable.

Pontécoulant follows Lagrange in the first edition of his *Théorie Analytique du Système du Monde* (1829), Tome II, Livre V, Ch. iv, No. 25, p. 395. He obtains, with different notation, the equation for ω derived from (17), viz.

$$\alpha_0 a^2 - \beta_0 b^2 = \frac{\omega^2}{2\pi\rho}(a^2 - b^2),$$

and then argues that because the right-hand member is independent of c the left-hand member must be so also; without regard to the

* Or Lagrange, *Œuvres*, Tome XI (1888), pp. 219, 464, where the above-mentioned errata are repeated.

Digitized by Google

fact that the equation is not an identity, but determines the value of ω. In note II (pp. 52–5) of his Supplement to Livre V, published subsequently to Jacobi's discovery, Pontécoulant supplies the developments of the theory necessitated by that discovery, which are incorporated in the text (pp. 430, 435–8) of the second volume of the second edition (1856) of his work, with a note (X) at the end (pp. 567–8) indicating the error in Lagrange's work; but with the repetition on p. 430 of the unsound argument just indicated.

In the *Roy. Soc. Phil. Trans.* for 1834, p. 513, Ivory gave another demonstration to the same effect as Lagrange, based on incorrect mechanics (see the paper by Todhunter on Jacobi's theorem and Ivory's discussion thereof in *Proc. Roy. Soc.*, 1870–1, vol. 19, pp. 42–56). Jacobi announced his discovery later in the same year (Poggendorff's *Annalen der Physik und Chemie*, Band 33, pp. 229–33; or Jacobi, *Werke*, Band II (1882), pp. 19–22), and in a further memoir published in the *Phil. Trans.* for 1838, pp. 57–66, Ivory proceeded to discuss the consequences. This memoir contains quite a number of unfounded assertions or inconclusive reasonings as to the signs of integrals, to some of which Liouville drew attention in 1839 in his *Journal de Mathématiques* (IV, pp. 169–174), whilst Todhunter has dealt with them more fully and completely in the paper cited above. In accepting a correction made by Liouville, Ivory (*Phil. Trans.* 1839, pp. 265–6) proceeds to make other unwarranted assertions, also dealt with by Todhunter.

It would appear that it was not till 1842 that the question was at all adequately treated, when C. O. Meyer, to whom the problem had been proposed by Jacobi himself, published in Crelle's *Journal* (XXIV, pp. 44–59) a demonstration that as the ratio of the greatest axis to the least increases to infinity, from the value corresponding to the Maclaurin-Jacobi junction, the ratio of the mean axis to the least constantly diminishes to unity (cf. § 5 above), whilst the angular velocity constantly diminishes to zero. In order to submerge refractory negative parts of his integrals Meyer ingeniously made use of an integral identity analogous to (8) above; a method which is substantially the same as that which lies at the foundation of Prof. Baker's paper (*l.c. ante*), and of Mr Hargreave's previous demonstrations in *Camb. Phil. Soc. Trans.* vol. 22, pp. 69–71. Unfortunately, in evaluating the above-mentioned identity, Meyer made the algebraical mistake of writing $4(s+t)$ for $3+s+t$; an error which vitiated his resulting expressions, though not his conclusions. For, on substituting the correct expression for the identity, precisely the same results follow, and even more directly.

Attention was drawn to Meyer's mistake, and correct expressions were given, by Liouville; who also extended the work on the same basis so as to shew that with a given mass, or volume, of liquid,

Digitized by Google

the moment of momentum increases as the ellipsoid elongates (*Additions à la Connaissance des Temps pour l'an* 1846, pp. 85–96; or *Journal de Mathématiques*, XVI, 1851, pp. 241–54). If with Meyer's notation we proceed to form the expression for the total energy (kinetic + potential), and then take the variation of this expression on the supposition that the volume and moment of momentum remain constant, we are led without difficulty to Jacobi's relation expressed in a form equivalent to (9), which the writer then found could be deduced from the usual form (7) by means of the integral identity (8). The second order differentiations, required to determine the question of stability, seem likely, however, to lead to expressions of some complexity by this method. It is curious that Todhunter does not appear to have been acquainted with Meyer's paper, although in his *History of the Mathematical Theories of Attraction*, etc. (Art. 1582) he refers to Bertrand's correction of Lagrange's error in his edition of the *Mécanique Analytique*, whilst at the end of his supplementary note (pp. 406–7) Bertrand makes mention of that paper. Had Todhunter been acquainted with Meyer's method it would seem to be doubtful whether he would have thought it worth while to devote so much attention to the development of the much less direct method of Ivory's analysis, in which the discussion of the problem turns on the range of values assumed by two quantities denoted by p and τ, equivalent in Prof. Baker's notation to

$$\sqrt{\frac{a-c}{c}}\cdot\sqrt{\frac{b-c}{c}} \text{ and } \sqrt{\frac{a-c}{c}}-\sqrt{\frac{b-c}{c}},$$

and therefore only somewhat indirectly related to the variations of the axes and of their ratios.

In a further paper published in *Proc. Roy. Soc.* 1872–3, vol. 21, pp. 119–21, Todhunter has drawn attention to another piece of erroneous reasoning in our subject, which appeared in 1866 in Poggendorff's *Annalen der Physik und Chemie*, Band 129, pp. 443–50, in a paper by Dahlander, purporting to show that the homogeneous rotating liquid ellipsoid can be in relative equilibrium when rotating about an axis which is not a principal axis. This, however, is a state which is easily shown to be impossible, as is done by Todhunter in the above cited paper. Dahlander's error proceeds in effect from the erroneous assumption that the "centrifugal force," relative to the fixed axis of rotation, can be analysed into three component "centrifugal forces" about three moving rectangular axes (proceeding from a point on the fixed axis); the component with reference to any axis being the same as it would be if the body were rotating about that axis as a fixed axis with angular velocity equal to its actual component angular velocity about the axis in question.

Digitized by Google

Lastly, in addition to the cases of inconclusive reasoning already cited, we have the two instances in Thomson and Tait's *Natural Philosophy* (2nd edition, Part II, 1883) dealt with in §§ 6 and 7 above. In their first edition of 1867 the authors confine themselves to proving that for any given values of a and b there is a value of c, without regard to the question whether that value is unique; and they do not deal specially with the Jacobian ellipsoid for which $a = b$ (§ 778′ of 2nd edition).

Digitized by Google

Periodic solutions of a linear differential equation of the second order with periodic coefficients. By Mr E. L. INCE, Trinity College.

[*Received* 14 November, *read* 7 December, 1925.]

The equation to be considered is of the type

$$\frac{d^2y}{dx^2} + p(x)\, y = 0,$$

where $p(x)$ is continuous for all real values of x, even, and periodic. It is no restriction to suppose that the period is π, and this assumption will be made, so that the equation is virtually Hill's equation.

Of two fundamental solutions the one is even and the other odd. In a previous paper, published by this Society*, it was proved that in a particular case, that of the Mathieu equation, the fundamental solutions cannot both be periodic and of period π or 2π. Other proofs of this theorem have subsequently been discovered, both by the present author and by other writers†. It is now proposed to examine the more general case‡.

By substituting in the equation a formal sine or cosine series-expression for the solution, it may be verified that if the solution is periodic, its period is an integer multiple of π. An even periodic solution will be denoted by $C(x)$, and an odd periodic solution by $S(x)$.

The expression for an even periodic solution, of period $s\pi$, is

$$C(x) = \sum_{k=-\infty}^{\infty} c_k \cos 2\left(\frac{r}{s} + k\right) x,$$

where r is prime to s, and $2r < s$ when $s > 2$.

Now the differential equation itself is unchanged if $\pi - x$ is written for x; if therefore it is satisfied by

$$y = C(x),$$

it is also satisfied by $\qquad y = C(\pi - x).$

But

$$C(\pi - x) = \cos\frac{2r\pi}{s}\, C(x) + \sin\frac{2r\pi}{s} \sum_{k=-\infty}^{\infty} c_k \sin 2\left(\frac{r}{s} + k\right) x.$$

When $s > 2$, neither $\cos\dfrac{2r\pi}{s}$ nor $\sin\dfrac{2r\pi}{s}$ is zero, and therefore when

* *Proc. Camb. Phil. Soc.* vol. 21 (1922), p. 117.

† E.g. Einar Hille, *Proc. Lond. Math. Soc.* (2), vol. 23 (1923), p. 224.

‡ Another particular case was discussed by the present writer, *Proc. Lond. Math. Soc.* (2), vol. 23 (1923), p. 56.

Digitized by Google

an even periodic solution exists, there exists also the odd periodic solution

$$S(x) = \sum_{k=-\infty}^{\infty} c_k \sin 2\left(\frac{r}{s} + k\right) x,$$

whose coefficients are precisely the same as those of the even solution.

When $s = 2$, $r = 1$ and the equation is satisfied by $y = C(x)$, then

$$C(\pi - x) = -C(x);$$

if it is satisfied by $y = S(x)$, then

$$S(\pi - x) = S(x).$$

When $s = 1$, $r = 1$ and the equation is satisfied by $y = C(x)$, then

$$C(\pi - x) = C(x);$$

if it is satisfied by $y = S(x)$, then

$$S(\pi - x) = S(x).$$

Now let $s = 2$ and consider the possibility of the coexistence of the even and odd periodic solutions

$$y = C(x), \quad y = S(x).$$

These solutions may be so chosen that

$$C(x) S'(x) - S(x) C'(x) = 1.$$

If

$$C(x) = \sum_{k=0}^{\infty} a_k \cos(2k+1) x,$$

$$S(x) = \sum_{k=0}^{\infty} b_k \sin(2k+1) x,$$

then

$$1 = \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} (2j+1) a_k b_j \cos(2k+1) x \cos(2j+1) x$$

$$+ \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} (2k+1) a_k b_j \sin(2k+1) x \cos(2j+1) x$$

$$= \sum_{k=0}^{\infty} (2k+1) a_k b_k + \sum_{(k>j)} \sum_{j=0}^{\infty} (k-j)(a_j b_k - a_k b_j) \cos 2(j+k+1) x$$

$$+ \sum_{(k>j)} \sum_{j=0}^{\infty} (j+k+1)(a_j b_k + a_k b_j) \cos 2(k-j) x.$$

This relation is satisfied identically, and therefore

$$\sum_{k=0}^{\infty} (2k+1) a_k b_k = 1,$$

$$\sum_{k=0}^{\infty} (2k+2)(a_k b_{k+1} + a_{k+1} b_k) = 0,$$

$$\sum_{k=0}^{\infty} (2k+3)(a_k b_{k+2} + a_{k+2} b_k) + (a_0 b_1 - a_1 b_0) = 0,$$

$$\sum_{k=0}^{\infty} (2k+4)(a_k b_{k+3} + a_{k+3} b_k) + 2(a_0 b_2 - a_2 b_0) = 0,$$

Digitized by Google

and so forth. Let the coefficients a_k be regarded as known, then there is an infinite set of linear equations to determine the coefficients b_k. But the formula

$$S(x) = C(x) \int \{C(x)\}^{-2}\, dx$$

shows that the coefficients b_k are uniquely determined (apart from a constant multiplier) by the knowledge of the coefficients a_k. Hence the system of linear equations has one and only one solution (b_k) such that the series-expression for $S(x)$ converges for all real values of x.

This solution is easily found; it is given by the relations

$$a_j b_k + a_k b_j = 0 \qquad (k - j \text{ odd}),$$
$$a_j b_k - a_k b_j = 0 \qquad (k - j \text{ even}),$$

and these relations are, in turn, satisfied by

$$b_k = (-1)^k a_k.$$

If, therefore, the equation

$$\frac{d^2y}{dx^2} + p(x)\, y = 0$$

has two periodic solutions, expressible as convergent infinite series, these solutions are of the forms

$$C(x) = \sum_{k=0}^{\infty} a_k \cos(2k+1)\, x,$$

$$S(x) = \sum_{k=0}^{\infty} (-1)^k a_k \sin(2k+1)\, x.$$

But now it is an easy matter to verify that $p(x)$ has the period $\frac{1}{2}\pi$, contrary to hypothesis. Consequently when $p(x)$ is of period π, the equation cannot admit of two solutions of period 2π.

If $C(x)$ terminates, and the last term is $a_m \cos(2m+1)\, x$, the second solution is periodic when, and only when,

$$a_0 = a_1 = \ldots = a_{m-1} = 0,$$

and then $$S(x) = b_m \sin(2m+1)\, x,$$

and $p(x)$ has the constant value $(2m+1)^2$.

It may also be proved that when $p(x)$ has the period π, two solutions, $C(x)$ and $S(x)$, of period π, cannot coexist except in the trivial case when $p(x)$ has the constant value $4m^2$. In these exceptional cases $p(x)$ is not, strictly speaking, of period π.

Digitized by Google

The second solution of the Mathieu equation. By Mr E. L. Ince, Trinity College.

[*Received* 14 November, *read* 7 December, 1925.]

The fact that the Mathieu equation, and, as was proved in the preceding paper, the more general Hill equation, can admit of but one solution of period π or 2π, suggests the problem of determining how far the second solution deviates from periodicity when the first solution has either the period π or the period 2π.

Let the Mathieu equation be written in the form

$$\frac{d^2y}{dx^2} + (a + 16q\cos 2x)\, y = 0,$$

then, if an even solution of period 2π exists, the *characteristic number* a must be a root of the determinantal equation

$$\begin{vmatrix} a-1+8q, & 8q, & 0, & \dots \\ 8q, & a-9, & 8q, & \dots \\ 0, & 8q, & a-25, & \dots \\ \dots & \dots & \dots & \dots \end{vmatrix} = 0.$$

The equation which determines the characteristic numbers for the odd solutions of period 2π is formed by changing the leading element of the above determinant into $a - 1 - 8q$.

Let a_{2n+1} be the characteristic number for the solution

$$y = ce_{2n+1}(x),$$

then a_{2n+1} is determined accurately to the term in q^{2n+1} if all but the first $2n + 1$ rows and columns of the determinant are ignored, and to a lower term q^{2m} $(m < n)$ if the determinant is stripped of all members except the m rows and columns on each side of the element $a - (2n + 1)^2$.

This shows that if

$$a_{2n+1} = (2n+1)^2 + \sum_{r=1}^{\infty} \alpha_r q^r,$$

then $$\alpha_1 = \alpha_3 = \dots = \alpha_{2n-1} = 0, \quad \alpha_{2n+1} \neq 0.$$

Incidentally it may be verified that, if $p = 2n + 1$,

$$\alpha_2 = \frac{32}{p^2 - 1} \qquad (p \geqslant 3),$$

$$\alpha_4 = \frac{128\,(5p^2 + 7)}{(p^2 - 1)^3\,(p^2 - 4)} \qquad (p \geqslant 5),$$

and so on.

Digitized by Google

Now let b_{2n+1} be the characteristic number for the solution

$$y = se_{2n+1}(x),$$

then, since $$b_{2n+1} = (2n+1)^2 + \sum_{r=1}^{\infty} (-1)^r \alpha_r q^r,$$

it follows that $$a_{2n+1} - b_{2n+1} = O(q^{2n+1}).$$

The corresponding theorem for solutions of period π is as follows. If a_{2n} and b_{2n} are the characteristic numbers for

$$ce_{2n}(x) \text{ and } se_{2n}(x)$$

respectively, a_{2n} and b_{2n} are even functions of q. If

$$a_{2n} = 4n^2 + \sum_{r=1}^{\infty} \alpha_{2r} q^{2r},$$

$$b_{2n} = 4n^2 + \sum_{r=1}^{\infty} \beta_{2r} q^{2r},$$

then $$\beta_{2r} = \alpha_{2r} \qquad (r < n),$$

$$\beta_{2n} \neq \alpha_{2n},$$

and consequently $$a_{2n} - b_{2n} = O(q^{2n}).$$

As before, if $p = 2n$,

$$\alpha_2 = \frac{32}{p^2 - 1} \qquad (p \geqslant 2),$$

$$\alpha_3 = \frac{128(5p^2+7)}{(p^2-1)^3(p^2-4)} \qquad (p \geqslant 4),$$

and so on.

The actual development of an even periodic solution in ascending powers of q is of the form

$$ce_p(x) = \cos px + \sum_{r=1}^{\infty} A_r q^r,$$

where

$$A_1 = \frac{2\cos(p+2)x}{p+1} - \frac{2\cos(p-2)x}{p-1} \qquad (p \geqslant 2),$$

$$A_2 = \frac{2\cos(p+4)x}{(p+1)(p+2)} + \frac{2\cos(p-4)x}{(p-1)(p-2)} \qquad (p \geqslant 3),$$

$$A_3 = \frac{4\cos(p+6)x}{3(p+1)(p+2)(p+3)} + \frac{4(p^2+4p+7)\cos(p+2)x}{(p-1)(p+1)^3(p+2)} - \frac{4(p^2-4p+7)\cos(p-2)x}{(p+1)(p-1)^3(p-2)} - \frac{4\cos(p-6)x}{3(p-1)(p-2)(p-3)} \qquad (p \geqslant 4),$$

Digitized by Google

and so on. Similarly, the odd solution $se_p(x)$ may be developed in the form

$$se_p(x) = \sin px + \sum_{r=1}^{\infty} B_r q^r,$$

where B_r is derived from A_r, for $r = 1, 2, \ldots, p-1$, by replacing each cosine by the corresponding sine.

Since b_p agrees with a_p up to, but not including, the term in q^p, the second solution corresponding to $ce_p(x)$ will agree with $se_p(x)$ up to and including the term in q^{p-1} at least.

If, therefore, the first solution is

$$y_1 = ce_p(x),$$

the second solution is of the form

$$y_2 = se_p(x) + \kappa x ce_p(x) + \phi(x),$$

where κ and $\phi(x)$ are of the order of q^p and κ is to be so chosen that $\phi(x)$ is odd and purely periodic.

Since y_2 is to satisfy the Mathieu equation for $a = a_p$, and $se_p(x)$ is a solution for $a = b_p$, it follows that $\phi(x)$ must satisfy

$$\phi''(x) + (a_p + 16q\cos 2x)\phi(x) + 2\kappa ce_p'(x) + (a_p - b_p)se_p(x) = 0.$$

Let $\kappa_0 q^p$ be the leading term of κ. Then $\phi(x)$ must satisfy

$$\phi''(x) + (a_p + 16q\cos 2x)\phi(x) = 2p\kappa_0 q^p \sin px - (a_p - \beta_p) q^p \sin px + O(q^{p+1}).$$

Since, therefore, $\phi(x)$ is to be purely periodic,

$$\kappa_0 = \frac{a_p - \beta_p}{2p},$$

and consequently $\phi(x) = O(q^{p+1})$.

Thus the periodic part of the second solution agrees with $se_p(x)$ up to and including the term in q^p. The non-periodic part begins with a term in q^p.

Digitized by Google

On a family of constructs in higher space. By C. G. F. JAMES, B.A., Trinity College.

[*Received* 24 November, *read* 7 December, 1925.]

The constructs in question are represented by the vanishing of all determinants of two rows and columns drawn from a matrix of $r+1$ rows and $s+1$ columns, where r can be taken less than or equal to s*. They are thus the multiple constructs of highest dimension on the well-known family of constructs generated by projective systems of linear spaces. In the present note the constructs are considered from a somewhat different point of view, the principal applications being to the determination of directrix loci of minimum order for the systems of spaces which they contain, and to the re-proof and extension of a known theorem.

§ 1. We can suppose all the constructs derived by section from those in a space, S, of $(r+1)(s+1)-1$ dimensions, given by

$$\begin{Vmatrix} x_1 & x_2 & \dots\dots & x_{s+1} \\ x_{s+2} & x_{s+3} & \dots\dots & x_{2s+2} \\ \dots & \dots & \dots & \dots \\ x_{rs+r+1} & \dots & \dots & x_{(r+1)(s+1)} \end{Vmatrix} = 0.$$

The coordinates in each row give the non-vanishing coordinates of $r+1$ linear spaces† $[s]$, any pair of which are in (1, 1) correspondence, corresponding points being such that the coordinates of the respective spaces in the same column of the matrix have the same value. The spaces $[r]$ joining corresponding points generate the construct which is thus of dimension $r+s$, and $f(r, s)$ denoting the order, will be denoted by $V_{r+s}^{f(r,s)}$. It similarly contains ∞^r spaces $[s]$. Such a construct and all its sections $V_{r+s-t}^{f(r,s)}$ will be said to belong to the same group. The work, in fact, illustrates the advantage of grouping together a set of constructs which can be obtained as general sections by spaces of suitable dimension of some non-conical construct of a higher space for which this is not true‡.

We now start *ab initio*, defining our constructs by means of the above collineation construction, and shall deduce therefrom such properties as are required even when such are already known by other methods.

* Cf. H. W. Richmond and F. Bath, *Proc. Camb. Phil. Soc.* vol. 22 (1924), § 6 ($r=1$).

† $[a]$ denotes a linear space of dimension a.

‡ For example the present $V_{r+s}^{f(r,s)}$, and examples quoted in "Complexes of conics and the Weddle surface," *Proc. Camb. Phil. Soc.* vol. 22 (1924), p. 5 (*small print*). Such constructs may be referred to as *section-normal*. In the paper quoted this property was proved only for r equal to unity, but from the present construction the extension to any value of r is immediate.

Digitized by Google

§ 2. Starting with the case $r = 1$ we obtain a $V_{s+1}^{f(1,s)}$ in $[2s + 1]$ generated by the lines, d, joining corresponding points in a collineation between two spaces S, S' of dimension s in independent position. It contains an infinite number of constructs $V_{j+1}^{f(1,j)}$ derived from spaces $[j]$ in the respective spaces S, S' and corresponding in the collineation. These we shall term subconstructs of $V_{s+1}^{f(1,s)}$.

(i) *The construct* $V_{s+1}^{f(1,s)}$ *contains* ∞^1 *spaces* $[s]$ *cut in collineation by the lines thereon.* (The latter statement follows immediately from the first.) Consider first the case $s = 2$, giving a V_3 in [5]. Let P be any point on V_3 and d the line through it. There are ∞^1 pairs of corresponding lines in S, S' meeting d and giving quadric subconstructs in spaces through d. The conjugate lines through P generate a cone C_2 which does not pass through d, since then S and S' would meet. Now consider the section of C_2 by a space of four dimensions S_4 through P. S_4 cuts S and S' in lines l, l' respectively, and l' cuts the line in S' corresponding in the collineation to l in a single point, thus giving a pair of corresponding points in S_4. Joining these to the intersections with d we have two corresponding lines defining a quadric subconstruct, such that the conjugate line through P lies in S_4, and conversely there is only one such line. C_2 is therefore a pencil, and through every point P of V_3 passes a plane contained therein. Let the theorem now be assumed true for $s = 2, 3, \ldots s - 1$. Take P as before on a line d of $V_{s+1}^{f(1,s)}$ and consider the section by an arbitrary space S_{2s} of dimension $2s$. This space cuts S and S' in $[s - 1]$ and as above it follows that S_{2s} contains in S, S' corresponding $[s - 2]$, which joined to the extremities of d give corresponding $[s - 1]$ in the respective spaces. The subconstruct defined by these $[s - 1]$ has by assumption an $[s - 1]$ through P. This cuts S_{2s} in an $[s - 2]$ (on the subconstruct defined by the corresponding $[s - 2]$ in S, S') and in P which is external to this space; and is thus contained therein. It forms the section by S_{2s} of the cone of conjugate lines, V_s, of the ∞^{s-1} quadric subconstructs arising from the ∞^{s-1} pairs of corresponding lines in S, S' meeting d. This cone is therefore linear and its lines form a space $[s]$. The theorem is thus proved generally.

The spaces $[s]$ *we shall regard as generating spaces, the lines* d *as directrices.*

(ii) *The order of the construct* is $s + 1$ $\{f(1, s) = s + 1\}$. Let us seek the number of intersections with a space $[s]$ in arbitrary situation. In the star of $[s + 1]$ passing through this space (which we shall denote by ϖ) a correspondence is established when the corresponding spaces cut S and S' in points corresponding in the

Digitized by Google

collineation between them. This is regarded as a correspondence of "points" in a space S_s, and we can apply a theorem of Schubert's*, which states that the number of united "points" equals the sum of the indices with a third number, which is now equal to the number of pairs of corresponding $[s+1]$ which lie in a $[2i]$ through ϖ and define an $[s+2]$ through a fixed $[s+1]$ in the $[2s]$. Now the $[2s]$ contains, as we have remarked above, corresponding $[s-2]$ in S and S' and the number sought is the number of joining lines of their collineation which meet an $[s-2]$, namely $f(1, s-2)$. The united "points" arise from the points where $V_{s+1}^{f(1,s)}$ meets ϖ. Hence

$$f(1, s) = 2 + f(1, s-2).$$

Now $f(1, 1) = 2$, and considering the above correspondence independently for $s = 2$ gives $f(1, 2) = 3$; and thus generally $f(1, s) = s + 1$. We pass on to the determination of directrix constructs for the generating spaces of the groups of constructs.

§ 3. *Directrix curves of the scrolls of the family.* These scrolls are in fact the well-known V_2^{r-1} of $[r]$, and from our point of view are obtained by taking the section of V_{s+1}^{s+1} of $[2s+1]$ by $[s+2]$. Previous to so doing, however, consider the subconstructs V_{j+1}^{j+1} of V_{s+1}^{s+1}. Their sections by $[j+1]$ in their spaces give normal curves of order $j+1$ and meeting each generating $[s]$ in a single point (in general). It is evident that such of these curves as lie in a $[2s+1-t]$ will be directrix curves of the $\infty^1 [s-t]$ on the V_{1+s-t}^{s+1} section of V_{s+1}^{s+1}. Now the dimension of the system of such curves equals the dimension of spaces $[j]$ in $[s]$ added to the dimension of the spaces $[j+1]$ in $[2j+1]$, namely†

$$(j+1)(s-j) + (j+2)j,$$

or

$$sj + s + j;$$

and the dimension of the system of those in a $[2s+1-t]$ will be

$$sj + s + j - t(j+2).$$

In the present case $(t = s-1)$ the dimension of the system of directrix curves C^{j+1} of the scroll V_2^{s+1} is in the general case $2j - s + 2$. In order then that such curves may really exist we must have $2j - s + 2 \geqslant 0$, and hence the least value j_0 of j compatible with this condition gives directrix curves of minimum order of this type on V_2^{s+1}. Though the existence of directrix curves of other type

* *Mat. Ann.* vol. 26, p. 56.

† It is readily seen that each $[j+1]$ arises once only in this construction.

Digitized by Google

is not excluded, yet such could not be of lower order. For if one such were of order $j_0 + 1 - h$ the scroll could be generated as the join of points of the two curves in (1, 1) correspondence and hence

$$s + 1 \leqslant j_0 + 1 + (j_0 + 1 - h)$$

which is impossible. We have thus re-proved the known* theorem:

A scroll V_2^{s+1} of $[s + 2]$ possesses in general a single directrix curve of minimum order $\frac{1}{2}s$ when s is even, and ∞^1 such curves of order $\frac{1}{2}(s + 1)$ when s is odd.

Scrolls with directrices of lower order can be obtained from V_{s+1}^{s+1} by taking the section by an $[s + 2]$ through such a directrix curve. Notice also that *the scroll of lines d* passing through points of a directrix curve C^{j+1} cuts each $[s]$ in a normal curve C^j. Conversely taking corresponding curves C^j in corresponding spaces $[j]$ in S and S' and considering the directrix curves C^{j+1} on the scroll of order $2j$ of lines joining corresponding points (the scroll being itself a member of the family we are considering) the same theory could be built up by induction. Since (Bertini, *loc. cit.*) all directrix curves on the scrolls are rational normal curves it follows inversely that they can all be obtained by one or other of these constructions, so that the above restriction is shewn to be superfluous.

§ 4. We shall now generalize these results to apply to the case of *any construct V_{s+1-t}^{s+1} section of V_{s+1}^{s+1} by a space* $[2s + 1 - t]$. By a previous argument the latter space contains corresponding $[s - 2t]$ in each space $[s]$, and hence *as long as*† $t \leqslant [\frac{1}{2}s]$ *the construct contains directrix lines d, and in general precisely ∞^{s-2t} thereof.* Curvilinear directrices are obtained as before by considering the sections of subconstructs derived from corresponding spaces $[j]$ in the $[s]$, and the condition for their existence on the general construct depends on the number j satisfying

$$sj + s + j - t(j + 2) \geqslant 0. \qquad \ldots\ldots(1)$$

To ensure that the least value of j (say j_0) satisfying this inequality does indeed give the minimum directrices we use a theorem of Schubert's‡ on series of ∞^1 groups of k points lying on a curve C (irreducible or compound) in a space $[d]$, the points of each group lying on an $[r]$ $(k > r)$. If x_0 denotes the number of groups of the series of which one point lies in a $[d - 1]$, x the number such that the joining $[r]$ meets a $[d - r - 1]$ in a point, y the number

* See Bertini, *Geometria proiettiva degli spazi...*, Ed. 2, Messina 1923, chap. 14, p. 538.

† $[x]$ as a numerical symbol denotes as usual the greatest integer contained in the positive number x.

‡ See Severi, "Sopra alcune singolarità delle curve...," *Mem. Torino* (2), vol. 51 (1902), p. 81.

Digitized by Google

of groups of which two (only) points coincide, and z the number, of whose points $r+1$ (only) belong to an $[r-1]$; then

$$y = 2x_0\frac{k-1}{r} - 2x\frac{k(k-1)}{r(r+1)} - 2z\Big/\binom{k-2}{r-1};$$

and hence
$$2x_0\frac{k-1}{r} - 2x\frac{k(k-1)}{r(r+1)} \geqslant 0.$$

Now take $s-t+1$ directrix curves and suppose (if possible) that all of these save one are of order j_0+1, the last being of lower order j_0+1-h. To this compound curve apply the last theorem, the groups being those on the generating $[s-t]$ of V^{s+1}_{s+1-t}. We have

$$k = s-t+1 = r+1, \quad x_0 = (s-t+1)(j_0+1) - h,$$

and
$$x = s+1.$$

Hence
$$(s-t+1)(j_0+1) - h - (s+1) \geqslant 0, \qquad \ldots\ldots(2)$$

so that
$$(s-t+1)j_0 \geqslant t+h.$$

Now, by (1), if j_0 is the actual minimum value as above

$$sj_0 + s + j_0 - t(j_0+2) \leqslant s-t,$$

or
$$(s-t+1)j_0 \leqslant t.$$

Comparing these two inequalities it is seen that the assumption was absurd. The argument excludes only the case when (1) becomes an equality when j_0 is written for j. In this case take the unique directrix curve C^{j_0+1}, $s-t-1$ directrix curves C^{j_0+2}, and another of order j_0+1-h (if such is possible). Then (1) gives

$$(s-t+1)j_0 = 2t-s,$$

and in place of (2) we have

$$(s-t+1)(j_0+1) + (s-t-1) - h - (s+1) \geqslant 0,$$

or
$$(s-t+1)j_0 \geqslant 2t-s+1+h,$$

whence
$$0 \geqslant h+1.$$

This is again impossible, and the value of j_0 actually gives the minimum directrices. Notice also that the last inequality shews further that in this case there can be only one directrix of minimum order. Thus

The construct V^{s+1}_{s+1-t} has ∞^a directrix curves of minimum order

$$\left[\frac{2t-s}{s+1-t}\right] + 2,$$

where a is defined as the residue when $2t-s$ is divided by $s+1-t$, provided it is not zero. In this excluded case there is a unique directrix curve of order

$$\left[\frac{2t-s}{s+1-t}\right] + 1.$$

Digitized by Google

Notice before passing on that the subconstructs obtained from corresponding $[s-1]$ in the $[s]$ give for every construct V_{s+1-t}^{s+1} a system of spaces $[2s-1-t]$ which cut it in V_{s-t}^{s}. In particular they give the s-secant $[s-1]$ of the rational normal curves C^{s+1} of the group, and the $[s]$ which cut the scroll V_2^{s+1} in rational normal s-tic curves ($t = s, s-1$ respectively).

§ 5. We pass on now to the general case in which the construct is $V_{r-s}^{f(r,s)}$ generated as the locus of spaces $[r]$ joining corresponding points in a set of collineations between $r+1$ spaces $[s]$. We verify first that in addition to the ∞^s spaces $[r]$ the construct also contains ∞^r spaces $[s]$. Let the original spaces $[s]$ be named $S_1, S_2, \ldots S_{r+1}$. Let P be a point on a line joining corresponding points in $S_1 S_2$. Then by the previous theorem through P passes an $[s]$ lying wholly in the V_{s+1}^{s+1} defined by S_1 and S_2 and therefore in the $V_{r+s}^{f(r,s)}$: and this $[s]$ is cut in collineation with $S_1 \ldots S_{r+1}$. Now let P' be any point on the line joining P to its correspondent in S_3. Through P' again passes an $[s]$ on the construct, and we join P' to its correspondent in S_4, selecting on the joining line a point P'', through which again passes an $[s]$ on the construct. It is evident that every point on the construct will be reached by this construction at or before the stage in which a point is selected on the line joining $P^{(r-2)}$ to its correspondent in S_{r+1}. Thus through every point passes an $[s]$ and the result follows.

To find the order of the construct notice first that in virtue of the theorem just proved $f(1,s) = f(s,1)$ (or generally $f(r,s) = f(s,r)$). If the $r+1$ spaces $[s]$ are chosen so as to lie in an $[r(s+1)]$ the construct so obtained is the projection of that of the general case and its order remains unaltered. Consider, therefore, the section of dimension $r+s-1$ by the $[rs+r-1]$, Σ, defined by (say) the first r of the $r+1$ spaces $[s]$. This consists of the $V_{r+s-1}^{f(r-1,s)}$ generated by the $[r-1]$ joining corresponding points in these r spaces $[s]$, together with the $V_{r+s}^{f(r,s-1)}$ generated by the $[r]$ joining corresponding points of $r+1$ spaces $[s-1]$ of which the last is the section of S_{r+1} by Σ and the remainder are its correspondents in $S_1 \ldots S_r$. Thus

$$f(r,s) = f(r-1,s) + f(r,s-1).$$

The solution of this equation subject to the above conditions is

$$f(r,s) = \binom{r+s}{r}$$

which is therefore the order of the construct.*

* This formula is not new, but the present proof I believe is. See C. Segre, "Gli ordine delle varietà...," *Atti Lincei* (5), vol. 9, ii (1900), p. 253.

Digitized by Google

§ 6. We now seek for directrices, in this case V_r, of the system of $\infty^r [s-t]$ in the section $V_{r+s-t}^{f(r,s)}$ of $V_{r+s}^{f(r,s)}$ by a space $[T]$ or $[(r+1)(s+1)-1-t]$. As before, provided t does not exceed a certain value, there will be directrix $[r]$. Now the $[T]$ contains corresponding $[s-t(r+1)]$ in $S_1 \ldots S_{r+1}$, and hence *there exist directrix* $[r]$ *provided*

$$t \leqslant \left[\frac{s}{r+1}\right],$$

the dimension of their system being $s-t(r+1)$.

It t exceeds this value we proceed as before to choose corresponding $[j]$ in the $[s]$, and to form the $V_{r+j}^{f(r,j)}$ in spaces $[jr+j+r]$ thereby defined. Their sections $V_r^{f(r,j)}$ by spaces $[jr+r]$ are directrix constructs for the system of $[s-t]$ on any section, provided their spaces lie in the $[T]$. The dimension of these directrices on the $V_{r+s-t}^{f(r,s)}$ is thus equal to the dimension of the system of $[j]$ in $[s]$ together with the dimension of the system of $[(j+1)r]$ in $[jr+j+r]$ less $\{(j+1)r+1\}t$, namely

$$(j+1)(s-j)+j(rj+r+1)-(rj+r+1)t.$$

The minimum value of j compatible with this condition gives the directrices of minimum order $f(r,j)$ of this type. The most interesting case arises when $t=s-1$, giving a construct of ∞^r lines. *The dimension of the system of directrix* $V_r^{f(r,j)}$ *is now*

$$(r-1)j^2-j(rs-2r-s)-(rs-r-1)$$

and the minimum value of j *is* θ, *or* $\theta+1$, *according as this expression can be made equal to zero or not, and the value of* θ *is*

$$\left[\frac{rs-2r-s+\{s^2(r-1)^2+4\}^{\frac{1}{2}}}{2(r-1)}\right].$$

The result in the present case, it must be pointed out, merely establishes the least value for the order of directrix constructs of a certain type. Whatever the results for $r=1$ may suggest, the possibility of the existence of other directrices of lower order is not excluded, as we have not devised an arithmetical test which shall exclude this possibility.

§ 7. *On the spaces of the subconstructs with reference to the points, lines, etc. of the fundamental space.* Taking first $r=1$ the spaces $[2j+1]$ of the subconstructs V_{j+1}^{j+1} form an interesting system Σ_{j+1} of dimension $(j+1)(s-j)$ $(j=1, 2, \ldots s-1)$. Such spaces we shall term $(j+1)$-spaces for V_{s+1}^{s+1}. Those passing through a point Q are obtained as follows. Through Q passes a line h

Digitized by Google

meeting the two spaces S, S' defining V_{s+1}^{s+1}, the points of section being (say) A, B' respectively. To B' corresponds in S a point B and to A a point A' in S'. The lines $AB, A'B'$ correspond and define an ordinary space meeting V_{s+1}^{s+1} in a quadric (namely a 2-space). The $(j+1)$-spaces $(j>1)$ through Q are obtained by taking corresponding $[j]$ through the lines $AB, A'B'$ respectively and hence their system is of dimension $(j-1)(s-j)$. Generally through a space $[k]$ in arbitrary position passes a space $[2k+1]$ cutting S and S' in $[k]$, so that there arise in these spaces corresponding $[2k+1]$, whose $[4k+3]$ passes through the original $[k]$, this space being a $(2k+2)$-space of V_{s+1}^{s+1}. The $(j+1)$-spaces $(j, s>2k+1)$ through the $[k]$ are obtained as before by drawing corresponding spaces $[j]$ through the $[2k+1]$ in S, S' and are therefore $\infty^{(s-2k-1)(s-j)}$ in number.

So also for arbitrary values of r there is one line through a point Q of the fundamental space meeting the defining spaces $S_1, S_2, \ldots S_{r+1}$ in points $A, B', C', \ldots$: to the latter r of these correspond in S_1 points $B, C, \ldots$. These points define a space $[r]$ therein (if $r<s$) to which correspond in $S_2, S_3, \ldots$ other spaces $[r]$ passing respectively through $B', C', \ldots$. The space $[r^2+2r]$ which they define passes through Q and is thus the unique $\binom{2r}{r}$-space passing through Q (to extend the above definition), cutting $V_{r+s}^{f(r,s)}$ in a $V_{2r}^{f(r,r)}$. As above there pass through Q $\infty^{(j-r)(s-j)}$ $\binom{r+j}{r}$-spaces $(j>r)$. Similarly through a $[k]$ passes one $\binom{rk+k+2r}{r}$-space, and $\infty^{(j-rk-k-r)(s-j)}$ $\binom{j+r}{r}$-spaces $(s, j>rk+r+k)$.

§ 8. The present method of discussing the constructs of our family leads to a new proof of a theorem on the 3-spaces of a V_4^4, namely*:

The ∞^2 3-spaces of a V_4^4 which meet a line l are such that those through a point of l pass through the same ordinary space (namely a 2-space), *whose locus as the point moves on l is a second V_4^4, any one of whose directrix lines may replace l. The spaces also pass by sets of ∞^1 through the generating spaces of a third V_4^4.*

These three constructs we shall denote by V_0, V_1, V_2, and we shall term V_1 the projecting construct and V_2 the derived construct of V_0 for the line l on V_1.

* The theorem as now worded summarises the results of my paper, "Extensions of a theorem of Segre's and their natural position in space of seven dimensions," *Proc. Camb. Phil. Soc.* vol. 21 (1923), pp. 664–84. See also *Id.* vol. 22, p. 24, for references.

Digitized by Google

The three constructs are in symmetrical relation in that starting with V_i its projecting and derived constructs with respect to a line of V_j are V_j and V_k respectively $(i \neq j \neq k)$. They cut in a rational sextic construct V_3^6, which the present method identifies with a construct obtained by an alternative method. The method has also the advantage of being capable of generalization, the new theorems being dealt with later.

Fixing two spaces S, S' of V_0 let Γ_0 denote the collineation between them as between each and all such spaces. Let P be a point on l, h the line through P meeting S in A and S' in B' (see § 8), and let A' in S' correspond to A in S, and B in S to B' in S'. Let m denote the line AA', m' the line BB'. (All lines named m belong to the system of directrices of V_0 and similarly all lines j, k eventually to those of V_2, V_1 respectively.) Then AB and $A'B'$ correspond in Γ_0.

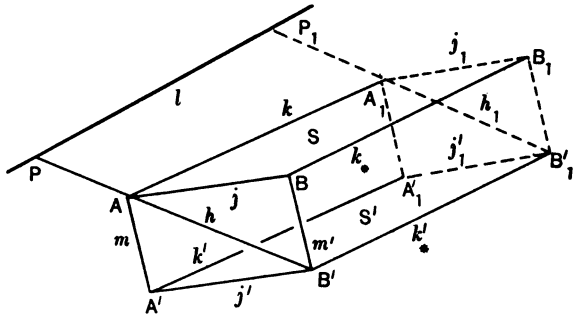


Fig. 1

As P varies on l, A varies on a line k, B' on a line k_*'. Let k' and k_* be their respective corresponding lines in S' and S. Then the lines j in the various positions join points on k and k_* which are thereby put into (1, 1) correspondence, and hence the lines j describe a regulus $R\,(jk)_j$, as we shall call it, of which k and k_* are directrices†. The lines j' describe the corresponding regulus $R'\,(j'k')_{j'}$, and there are corresponding figures in all the spaces (S) on V_0. The lines analogous to j or j' in these describe a regulus $R\,(jm)_j$ of which m and m' are directrices, and so also do the sets of lines analogously derived from any point on l. Let P_1 denote a second point on l. Let the analogous figure be constructed, all its points, lines, etc. being qualified by a lower suffix 1 (see Fig. 1).

Now all the 3-spaces through P pass through the spaces (jj') and so also those through any point on l pass through the space

† Generally $R\,(ab)_a$ denotes a regulus whose lines belong to a system each line of which is named "a," and the directrices belong to a system "b," the various symbols R, a and b being possibly qualified by accents, suffices, etc. $Q\,(ab)$ denotes the quadric base of these reguli.

Digitized by Google

defined by the lines j analogously derived. These spaces contain quadrics $Q(jm)$, etc., bases of the reguli $R(jm)_j$ in their various positions; and these quadrics are in collineation in pairs, all corresponding points lying on a line. This is therefore subordinate to a collineation Γ_1 between their spaces, which therefore generate a second V_4^4 namely V_1, the projecting construct. To this l belongs as directrix, and may therefore be replaced by any other directrix, all of which (except l) are named "k."

We have seen that all the 3-spaces meeting l are obtained from planes passing through one or other of the lines j in S†. Each therefore passes through a line of the regulus $R(jk)_k$ conjugate to $R(jk)_j$, and thus the spaces in question pass by sets of ∞^1 through a second set of ordinary spaces, namely those joining lines of this regulus to their corresponding lines in S'. The quadrics $Q(km)$ in these spaces (subconstructs of V_0) are again in collineation as above, and their spaces therefore describe a third V_4^4 namely V_2, the derived construct. The collineation of this construct we term Γ_2 and its directrices are named "j."

The three constructs cut in the locus of the quadrics $Q(jk)$, $Q(jm)$, and $Q(km)$, which is thus triply described as the join of quadrics in collineation, and is thus identical with a construct mentioned by Dr Richmond and Mr Bath‡. It is a sextic, V_3^6, in virtue of the lemma:

The construct which is generated by the joins of corresponding points of two V_d^R *in collineation* (namely which correspond in a fixed collineation between their spaces) *is of order* $R(d+1)$. *More generally given* $r+1$ *such constructs in collineation the order of the locus of spaces* $[r]$ *joining corresponding points is* $R\binom{d+r}{r}$.

The lemma is deduced at once by an appeal to the conservation principle. To cover the results quoted above, it remains to prove the reciprocity of the various quartic constructs. The reciprocal relation of V_0 and V_2 for V_1 as projecting construct is the most quickly proved since we can take the same line l and consider the 3-spaces of V_2 which meet it. Let S_2 joining k and k', and S_2' joining k_*, k_*' be assumed as the defining spaces of V_2: then the line through P meeting these spaces is again h, but m and m' are now the lines through A, B' respectively and corresponding in Γ_2. As P varies A, A', B, B' move as before on k, k', k_*, and k_*'; and the locus of the space (mm'), the new projecting construct, is again V_1.

† The system is thus represented on the tangent planes of a quadric, and therefore also on the points of a quadric such that those through a point are represented by points on a line of one regulus. This result, read off from the equation of the construct, was the starting-point of my other method of proof. "Extensions, etc." *loc. cit.*

‡ *Loc. cit.*

Digitized by Google

The derived construct is the locus of (jj') in its various new positions, namely V_0.

We pass on to prove that starting with a line of V_0 and considering the 3-spaces of V_1 we obtain V_0 as projecting construct and V_2 as derived construct. It is important to remark that *the points, lines, etc. which we proceed to define are not in the first place to be identified with the points, lines, etc. which bear the same names in the previous figure, with the exception of the lines* j, j', j_1 *and* j_1'. Let S_1 joining jj' (Fig. 2) and S_1' joining $j_1 j_1'$ be assumed as the spaces

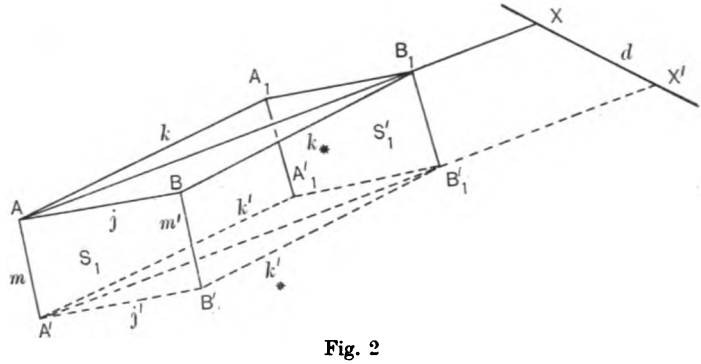


Fig. 2

defining V_1. Let d (a line "m") be a directrix of V_0 joining corresponding points X, X' in S, S'. Let g be drawn in S so as to meet j in A and j_1 in B_1 and passing through X. This is the line (analogous to h) which meets both defining spaces (now) of V_1. The lines of these spaces passing through A and B_1 and corresponding in Γ_1 are j and j_1. In this same correspondence let A_1 on j_1 correspond to A and B on j to B_1. Also let A', B', A_1', B_1' of S' correspond to A, B, A_1, B_1 in Γ_0 as in the former figure, and let the joining lines be named as in the former figure when they join similarly named points.

As X varies on d there arise similar figures in the generating spaces on V_0, and in particular in the space S', each figure corresponding in Γ_0 to that in S or S'. A varies on m and B on m'. The projecting construct is the locus of the spaces analogous to (jj_1) in the various figures, namely V_0.

The derived construct is the locus of the space joining a line of the regulus $R(jm)_m$ to its correspondent in the collineation between S_1 and S_1', and this is seen to be identical with V_2. These two cases cover the full reciprocal relations of the three constructs.

§ 9. *Theorems analogous to the theorem of* § 8. The new method of proof of the theorem we have been dealing with shews that it is one of a large number of similar theorems dealing with constructs of the family we are considering. We shall restrict ourselves to the

Digitized by Google

case when these constructs are in the highest space possible. Each gives rise to a group of theorems, respectively for the various sections of the section-normal construct*, but we shall not, as a rule, record these individually. We deal first with the most completely analogous results. Let $r+1$ spaces $[s]$ be taken in collineation and the $V_{r-s}^{f(r,s)}$ constructed, where s for the moment is undetermined. Through a point P external to the construct passes a line h meeting the chosen spaces $S_1, S_2, S_3, \ldots S_{r+1}$ in points $A, B', C'', \ldots$ etc. Then to $B', C'', \ldots$ etc. correspond in S_1 points $B, C, \ldots$; and similarly to $A, C'', D''', \ldots$ etc. points $A', C', D', \ldots$ etc. in S_2, and so on; and hence there arise in each space corresponding $[r]$, passing respectively through $A, B', C'', \ldots$ etc. and therefore such that the $\binom{2r}{r}$-space (of dimension r^2+2r) they define passes through P.

As P varies on a space $[r]$ external to the construct, which we shall denote by ϖ, the points $A, B, C, \ldots$ vary on spaces $[r]$ in S_1 and the spaces $[r]$ they define in their various positions join corresponding points in a collineation Γ_2, and therefore describe a $V_{2r}^{f(r,r)}$ in S_1, and there are collinear constructs in $S_2 \ldots S_{r+1}$. We shall suppose that these are normal and belong to the $[s]$ so that $s = r^2+2r$: other cases are of minor interest and are noted later. Until further stated therefore s is to be understood as being written for the sake of brevity in place of this expression. This system of spaces $[r]$ in $V_{2r}^{f(r,r)}$ will be denoted by Σ, the conjugate system, containing the loci of A, B, etc. by Σ'. Now an $[s-1]$ through a space of Σ meets $V_{2r}^{f(r,r)}$ in a space of Σ', for the space, η, of Σ chosen, may be taken with r others to define the system Σ', and these r others are cut by the $[s-1]$ in $[r-1]$, and hence there arise in these spaces corresponding points (uniquely determined in each) which joined to their correspondents in η give a space of Σ' in the $[s-1]$. Now such $[s-1]$ all meet ϖ, and define with their correspondents in S_2, etc. $\binom{r+s-1}{r}$-spaces of $V_{r+s}^{f(r,s)}$. In future we shall say that a complete set of $r+1$ corresponding figures in $S_1 \ldots S_{r+1}$ correspond in Γ. The spaces defined by each space of Σ and its correspondents in Γ, which are also spaces of dimension s, correspond in a collineation Γ_1, determined by the collineation between the spaces of Σ given by the spaces of the conjugate system Σ'. They therefore define a second $V_{r+s}^{f(r,s)}$ and the spaces of Σ' give a third such construct in the same way. The reciprocal relation between these three constructs is proved by the method used in the special case $r=1$, and the following theorem is thus obtained. In it s denotes r^2+2r.

* Cf. James, *loc. cit.*

Digitized by Google

Theorem I. The $\infty^{s-1}\binom{r+s-1}{r}$*-spaces of dimension* $rs+s-1$ *of a* $V_{r+s}^{f(r,s)}$ *in* $[rs+r+s]$ *meeting an external space* ϖ *of the* r*th dimension pass by* ∞^{s-1-r} *through the generating spaces of an analogous construct,* the projecting construct, *these generating spaces being the* $\binom{2r}{r}$*-spaces of the original construct through points of* ϖ*. The first-named spaces also pass by* ∞^{s-1-r} *through the generating spaces of a third such construct,* the derived construct, *these spaces being also* $\binom{2r}{r}$*-spaces of the original construct.*

The three $V_{r+s}^{f(r,s)}$ *are in symmetrical relation* (as explained for the special case $r=1$, § 8) *and cut in a construct of dimension* $3r$ *and order** $(3r)!/(r!)^3$, *described either in three ways as locus of* ∞^r V_{2r}^{x} $\left\{x=\binom{2r}{r}\right\}$ *corresponding in collineations between spaces in certain respective systems* ∞^r*, or in three ways as the locus of* ∞^r *spaces joining corresponding points in collineations between* $r+1$ *such* V_{2r}^{x}.

In this case only we shall state the theorem given by cutting the figure with the least space possible to give a significant result, namely:

Theorem I (a). The ∞^{s-1} *spaces* $[r^3+2r^2-r]$ *meeting an* $[r]$ *and* $\binom{r^2+3r-1}{r}$*-secant a certain curve of order* $\binom{r^2+3r}{r}$ *in an* $[r^3+2r^2+1]$ *all meet a second such curve. The two curves cut in* $(3r)!/(r!)^3$ *points.*

Alternatively we have in such a space *sets of three such curves cutting in* $(3r)!/(r!)^3$ *points, common to each, and such that any space* $[r^3+2r^2-r]$ *meeting one and* $\binom{r^2+3r-1}{r}$*-secant a second also meets the third such curve.*

The curve in question is given analytically by its equation (see the introduction).

A note on the case $s>r^2+2r$. To illustrate what happens in this case† we may select for brevity the case $r=1$, and refer to the notation, etc. of § 8. As P varies on the line l the quadric in S is described by j as before, and we are now to consider spaces $[s-1]$ passing through a line j, and defining with their correspondents in Γ s-spaces of V_{s+1}^{s+1}; and thus those s-spaces which meet a line l pass through the generating spaces [3] of a V_4^4 (these spaces being the 2-spaces through points of l) and through those of a second V_4^4.

* By the lemma, § 8, namely in the first place $\binom{r+2r}{r}\binom{2r}{r}$.

† Cf. James, "Extensions...," *loc. cit.* § 15.

Digitized by Google

This result is not, however, of any interest as it deals essentially with the construct in the unique 4-space of V_{s+1}^{s+1} which passes through l, and whose space is in fact cut by the s-spaces in 3-spaces (of V_4^4 and V_{s+1}^{s+1}).

The case $s < r^2 + 2r$ is however more significant. Suppose in fact

$$s = r^2 + 2r - i,$$

where i must not be so large that two $[r]$ in an $[s]$ cut, namely i must not exceed $r^2 - 1$. The theorem obtained refers as before to $\binom{r+s-1}{r}$-spaces of a $V_{r+s}^{f(r,s)}$ in $[rs + r + s]$ meeting an external space $[r]$, but the projecting and derived constructs are no longer of the same type but are $V_{r+s+i}^{f(r,s+i)}$, and since these constructs are not normal the reciprocal properties no longer hold. The construct common to the original projecting and derived construct has the same dimension and order but is no longer normal.

In the remaining discussion we shall not refer to the cases analogous to those just considered.

§ 10. Theorem I is not the only result deducible from the figure of § 8. For let us consider a space $[x]$ in S_1 passing through $c + 1$ of the spaces $[r]$ of the system Σ, and let these be regarded as defining spaces of the $V_{2r}^{f(r,r)}$ in S_1. Then the $[x]$ cuts the remaining $r - c$ defining spaces in $[r + x - s]$, and we obtain by the usual argument corresponding $[r - (r - c)(s - x)]$ in these spaces. If these are points only, or if $r - (r - c)(s - x) = 0$, the space $[x]$ contains a space of the conjugate system. The two following cases always arise:

$$c = 0, \qquad x = s - 1;$$

and

$$c = r - 1, \quad x = s - r;$$

of which the former gives *Theorem I*. The spaces $[x]$ in the latter case joined to their correspondents in Γ give $\binom{r^2+2r}{r}$-spaces of $V_{r+s}^{f(r,s)}$ meeting ϖ in spaces $[r - 1]$. The number of spaces $[x]$ passing through one of dimension t in $[s]$ is ∞^a, where

$$a = (s - x)(x - t),$$

namely in the present case $a = r$. To see the nature of the projecting construct consider first the locus of the spaces $[r^2 + 2r]$ joining the spaces $[r]$ of Σ to their correspondents in Γ, namely the projecting construct, a $V_{r+s}^{f(r,s)}$ $(s = r^2 + 2r)$. The actual space common to all the spaces $[xr + x + r]$ through an $[r - 1]$ in ϖ is obtained as the join of r of the generating spaces $[s]$ of this construct, and is therefore of dimension $r^3 + r^2 + r - 1$, and can

Digitized by Google

be alternatively obtained by joining corresponding $[r-1]$ in $s+1$ collinear $[r]$. It can be readily verified that the locus of these spaces is the* $V_{r^3+2r^2+2r-1}$, represented by the vanishing of determinants of $r+1$ rows and columns from the matrix in § 1. Generally in fact:

The locus of spaces $[rc+r+c]$ *joining corresponding spaces* $[c]$ *in* $s+1$ *collinear spaces* $[r]$ *in space of dimension* $rs+r+s$ *is the construct of dimension* $(r+s)(c+1)-c^2$ *represented analytically by the vanishing of determinants of* $c+2$ *rows and columns drawn from the fundamental matrix of* § 1. From this it is seen that it must also be the locus of joins of spaces $[c]$ corresponding in the collineation between $r+1$ of the spaces $[s]$ of the conjugate system. For the order of the construct, which we shall denote by $f(r, s: c)$, see Segre, *loc. cit.* Constructs which are derived from the same spaces in collineation for various values of c we shall say are associated.

Applying this to the present case we have in all the theorem:

Theorem II. The $\infty^{2r}\binom{r^2+2r}{r}$*-spaces of the construct of Theorem I, being spaces of dimension* r^3+2r^2+2r, *meeting the external space* ϖ *in* $[r-1]$ *pass by* ∞^r *through the* $\binom{r^2+2r-1}{r}$*-spaces of the construct which meet* ϖ *singly in these spaces* $[r-1]$, *and which generate the* $V^{f(r,s:r-1)}_{r^3+2r^2+2r-1}$ *associated with the projecting construct. The* $\binom{r^2+2r}{r}$*-spaces in question also pass by* ∞^r *through the generating spaces* $[s]$ *of the derived construct, both projecting and derived constructs being the same as in Theorem I.*

§ 11. In special cases there are other theorems deducible from the same figure. Let r have any two complementary factors p, q $(p, q > 1)$, so that we can take

$$c = r - p = p(q-1) : x = s - q = p^2q^2 + 2pq - q,$$

and using c, x, etc. in the enunciation for the sake of brevity the new theorem is:

Theorem II (a). The $\infty^{r^2+r-q^2+p(c+1)}\binom{x+r}{r}$*-spaces of the original construct* $V^{f(r,s)}_{r+s}$ $(s = r^2+2r)$ *which pass through spaces* $[r-p]$ *in* ϖ *pass by* $\infty^{r^2+r-q^2}$ *through the* $\binom{cr+c+r}{r}$*-spaces of the same construct meeting* ϖ *singly in the same spaces* $[r-p]$, *and generating the* $V^{f(r,s:r-p)}_{r^3+3r^2+3r-p(r^2+r+p)}$ *associated with the projecting construct. They also pass by* $\infty^{r^2-q^2+p(c+1)}$ *through the generating spaces* $[s]$ *of the derived construct.*

* Of order $\binom{s+1}{r}$ namely here $\binom{r^2+2r+1}{r}$. See Segre, *loc. cit.*

Digitized by Google

Other theorems are obtained by removing the condition that the spaces $[x]$ shall contain only one space of the system Σ'.

§ 12. Let us now return for a moment to the case $r = 1$, and let there be defined a V_{s+1}^{s+1} by means of collinear spaces $[s]$ in $[2s + 1]$. Through an external point P passes one line h to meet these spaces, S and S' in A, B' (cf. § 8); and hence arise in them corresponding lines AB, $A'B'$, or (say) j, j', whose space passes through P. As P varies on an external space ϖ of dimension t, the points A and B defining j vary on spaces $[t]$ in S, and hence j on the directrix system of a V_{t+1}^{t+1} of $[2t + 1]$, and there is a corresponding construct in S'. Taking then $s = 2t + 1$ we build up the figure as in previous cases.

Taking $t + 1$ rays j to define the system of generating spaces $[t]$ on the V_{t+1}^{t+1} in S it is evident that a space $[s - 1]$ or $[2t]$ through t of these rays (and therefore one through any t rays of the system) contains one generating space $[t]$. Using this the following theorem is obtained as before:

Theorem III. The ∞^{t+1} $(2t + 1)$*-spaces of dimension* $4t + 1$ *of a* V_{2t+2}^{2t+2} *in* $[4t + 3]$ *which pass through* $[t - 1]$ *of a fixed external space* ϖ *of dimension* t *pass by* ∞^1 *through the* $2t$*-spaces (of dimension* $4t - 1$*) through the single spaces* $[t - 1]$ *in question, and again by* ∞^1 *through the spaces* $[2t + 1]$ *of a second* V_{2t+2}^{2t+2}*, say the derived construct.*

The locus of the $2t$*-spaces in question is the* $V_{5t-1}^{f(3,t:t-1)}$ *associated with the* $V_{t+3}^{f(3,t)}$ *locus of the ordinary spaces joining the lines* j *to their correspondents in* S'*. There is in consequence reciprocity between the initial and derived constructs in that each is obtained from the other, selecting as external space* ϖ *in the construction a space on the* $V_{t+3}^{f(3,t)}$*. The three constructs in question cut in a* $V_{t+2}^{(t+1)(t+2)}$*, locus (in two ways) of* ∞^1 V_{t+1}^{t+1} *in collineation.*

This theorem and Theorem I are alternative extensions of the initial theorem of § 8. They and the other theorems quoted are included in the following final case in which there are $r + 1$ spaces $S_1, S_2, \ldots$ of dimension s in collineation, while the typical external point P varies on an external space ϖ of dimension t. In each space $[s]$, in particular in S_1, we have a construct $V_{r+t}^{f(r,t)}$ containing a system Σ of spaces $[r]$ and a system Σ' of spaces $[t]$. The spaces Σ are such that joined to their correspondents in Γ they define spaces $[r^2 + 2r]$ passing through single points P in ϖ and which are $\binom{2r}{r}$-spaces of the fundamental construct derived from the spaces

Digitized by Google

$[s]$. The spaces of Σ' are the conjugate spaces. We are to take $s = rt + r + t$, the fundamental construct being therefore a $V^{f(r,\,rt+r+t)}_{rt+2r+t}$, say V_0.

Consider now the space defined by $c + 1$ spaces $[r]$ of Σ, a space of dimension $cr + c + r$. Joined to its corresponding spaces in Γ it defines a $\binom{cr+c+r}{r}$-space of V_0 meeting ϖ in a space $[c]$, and any space $[x]$ through it will define an $\binom{xr+x+r}{r}$-space with the same property. Let the $c + 1$ spaces be chosen from $t + 1$ spaces $[r]$ of Σ forming a set to define the system Σ' of spaces $[t]$ on the $V^{f(r,\,t)}_{r+t}$ in S_1. The remaining $t - c$ spaces $[r]$ are cut by the $[x]$ in $[r - (c - x)]$, and therefore as before contain $[r - (t - c)(s - x)]$, corresponding in the collineation between the $[r]$ of Σ. If this expression be zero, namely, if we have corresponding points only, then these points joined to their correspondents in the first $c + 1$ spaces define a space $[t]$ of Σ' in the $[x]$ in question. The subsequent details are built up as before. The cases which universally arise are

$$\text{(i)}\quad x = s - 1, \quad c = t - r,$$

$$\text{(ii)}\quad x = s - r, \quad c = t - 1,$$

but if r have any two complementary factors p, q we can take

$$\text{(iii)}\quad x = s - q, \quad c = t - p;$$

and we shall state the theorems simultaneously using the symbols s, x, c, etc. to denote these various values. Further, in the enunciation α denotes $r^2 + r - q^2$, and β denotes $p(c + 1)$ or $p(t - p + 1)$. The theorem obtained is:

Theorem IV. The $\infty^{\alpha+\beta}\binom{x+r}{r}$*-spaces of dimension* $xr + x + r$ *of the* $V^{f(r,\,s)}_{r+s}$ *subsequently denoted by* V_0 *in* $[rs + r + s]$ $(s = rt + r + t)$ *which meet an external space* ϖ *of dimension* t *in* $[t - p]$ *pass by* ∞^{α} *through the* $\binom{cr+c+2r}{r}$*-spaces of dimension* $(cr + c + r)r + cr + c + 2r$ *of* V_0 *passing through the single* $[t - p]$ *in* ϖ *and by* $\infty^{\alpha+3-r}$ *through the generating spaces* $[s]$ *of a second* $V^{f(r,\,s)}_{r+s}$*, say* V_2*, the derived construct.*

The locus of the $\binom{cr + c + 2r}{r}$*-spaces in question is the* $V^{f(r^2+2r,\,t:\,c)}_{(r^2+2r+t)(c+1)-c^2}$ *associated with the* $V^{f(r^2+2r,\,t)}_{r^2+2r+t}$ *locus of the* $[r^2 + 2r]$ *joining spaces of* Σ *in* S_1 *to their correspondents in* Γ.

The relation of V_0 *and* V_2 *is reciprocal in that each is derived from*

Digitized by Google

the other taking any space $[t]$ *of the* $V^{f(r^2+2r,t)}_{r^2+2r+t}$ *as the space* ϖ *of the construction, but there does not appear to be any reciprocal relation between* V_0 *or* V_2 *and this last-named construct, save in the case of Theorem I. The three constructs cut in a* V_{2r+t}, *locus (in two ways) of* $\infty^r V^{f(r,t)}_{r+t}$, *and of order* $\binom{r+t}{r}\binom{2r+t}{r}$.

Digitized by Google

The equianharmonic and harmonic envelopes of a quadrinodal cubic surface. By G. TIMMS, St John's College. (Communicated by Mr W. P. MILNE.)

[*Received* 31 October, *read* 7 December, 1925.]

Prof. W. P. Milne has recently been investigating, in connection with the general cubic surface, the Equianharmonic Envelope S_4 and the Harmonic Envelope T_6, defined as the envelopes of planes cutting the surface in equianharmonic and harmonic cubic curves respectively: he suggested to me that I should consider the particular case of the quadrinodal cubic surface.

Denoting the quadrinodal cubic surface $\Sigma \frac{1}{x} = 0$ by U_3, and the unique quadric having as enveloping cones the nodal cones of U_3 by Φ_2 ($\equiv \Sigma l^2 - 2\Sigma lm$), we obtain by direct calculation:

The tangential equation to U_3 is $\quad \Phi_2^2 - 64lmnp = 0.$

The Equianharmonic Envelope S_4 is $\Phi_2^2 - 48lmnp = 0.$

The Harmonic Envelope T_6 is $\quad \Phi_2(\Phi_2^2 - 72lmnp) = 0.$

Hence:

The Equianharmonic Envelope S_4 of the quadrinodal cubic surface is a surface of the fourth class, having as nodal points and nodal cones the nodal points and nodal cones of the cubic surface.

The Harmonic Envelope T_6 of the quadrinodal cubic surface consists of the quadric having as enveloping cones the nodal cones of the cubic surface and a surface of the fourth class having the same nodal points and nodal cones as the cubic surface.

The degeneration of T_6 into a quadric and a quartic envelope is easily seen geometrically, for the enveloping cone drawn from any point P on U_3 to U_3 consists of two quadric cones, and the envelope of a plane through P cutting these two cones in pairs of lines which harmonically separate each other will be the enveloping cone drawn from P to the quadric Φ_2.

That S_4 and the quartic part of T_6 have the same nodal cones may be seen by considering the fact that a nodal cubic curve of finite cross-ratio is cuspidal.

Digitized by Google

The adiabatic hypothesis for magnetic fields. By P. A. M. DIRAC, St John's College. (Communicated by Mr R. H. FOWLER.)

[*Received* 5 November, *read* 7 December, 1925.]

The validity of Hamilton's equations during an adiabatic change.

The usual proof of the adiabatic invariance of the quantum integrals of a multiply periodic dynamical system rests on the assumption that the Hamiltonian equations remain valid for the adiabatically changing system, with the same Hamiltonian function H, or at any rate with a new Hamiltonian function differing from the old one only by terms that vanish with $\dot{a}$, (a) being the varying parameter*.

Now in most cases of an adiabatic change, including all cases when a mechanical constraint of the system is adiabatically changed, the accelerations $(\ddot{x}, \ddot{y}, \ddot{z})$ of the various particles composing the system are the same functions of the co-ordinates (x, y, z) and velocities $(\dot{x}, \dot{y}, \dot{z})$ as for the undisturbed motion. For these cases, if the relations connecting the Hamiltonian p's and q's with the (x, y, z)'s and $(\dot{x}, \dot{y}, \dot{z})$'s do not contain the varying parameter (a) explicitly, the $\dot{p}$'s and $\dot{q}$'s must be the same functions of the p's and q's as for the undisturbed motion, so that the Hamiltonian equations will still be satisfied. If the p's and q's do involve the parameter (a) explicitly, we should have instead

$$\left.\begin{aligned} \dot{p}_k &= -\frac{\partial H}{\partial q_k} + \frac{\partial p_k}{\partial a}\dot{a} \\ \dot{q}_k &= \frac{\partial H}{\partial p_k} + \frac{\partial q_k}{\partial a}\dot{a} \end{aligned}\right\} \quad \text{.....(i),}$$

where the p_k and q_k are regarded as functions of the (x, y, z)'s, $(\dot{x}, \dot{y}, \dot{z})$'s and (a) in the partial differential coefficients with respect to (a).

Whenever there are no forces in the system depending on the velocities, it is always possible to choose a set of p's and q's that do not contain (a) explicitly; we can, in fact, take the Cartesian co-ordinates (x, y, z) as q's, when the corresponding p's will be $m\dot{x}$, $m\dot{y}$, $m\dot{z}$. If there is one set of variables p_k, q_k not containing (a) explicitly, any other set p_k', q_k' involving (a), defined by

$$\Sigma p_k dq_k = \Sigma p_k' dq_k' + dS,$$

will still satisfy the Hamiltonian equations, though with a new

* Dirac, *Proc. Roy. Soc.* A, vol. 107 (1925), p. 726.

Digitized by Google

Hamiltonian function H' differing from the original one by $\partial S/\partial a \,.\, \dot{a}$, since the relation

$$\Sigma p_k dq_k - H\,dt = \Sigma p_k' dq_k' - H'\,dt + dS$$

will still be true with (a) varying.

For a system which does involve forces depending on the velocities, there may be no set of p's and q's satisfying the Hamiltonian equations during the adiabatic motion, in which case the quantum integrals will not, in general, be constant.

Effect of an adiabatically varying magnetic field on an atomic system.

The only important practical case of a system which does involve forces depending on the velocities is that of an atom in an external magnetic field. The Lagrangian of such a system is given by

$$L = L_0 - e/c.\Sigma\,(A_x\dot{x} + A_y\dot{y} + A_z\dot{z}),$$

where L_0 is the Lagrangian of the system in the absence of the magnetic field, (A_x, A_y, A_z) is the vector potential of the field, and the summation is extended over all the electrons of the system. The momenta are given by equations of the type

$$p_x = \frac{\partial L}{\partial \dot{x}} = m\dot{x} - \frac{e}{c}A_x, \qquad \text{.....(ii)}$$

and Hamilton's equations are satisfied, to the first order in $\boldsymbol{A}$, with the Hamiltonian function

$$H = H_0 + \Sigma\, e/c.(A, v)^*,$$

H_0 being the Hamiltonian in the absence of the field, and v being the velocity of an electron.

For a uniform magnetic field $\mathbf{H}$ we may take

$$A = \tfrac{1}{2}\,[\mathbf{H}, r], \qquad \text{.....(iii)}$$

where r is the position vector of the electron considered, this being one of the solutions of the equation

$$\mathbf{H} = \text{curl}\, A.$$

This gives

$$H = H_0 + e/2mc.|\,\mathbf{H}\,|\,p_\phi, \qquad \text{.....(iv)}$$

where p_ϕ is the component of the resultant angular momentum of the system in the direction of the field. Since the new term $e/2mc.|\,\mathbf{H}\,|\,p_\phi$ added to the Hamiltonian remains constant during the motion, the quantum integrals are given by the same functions of the p's and q's as in the absence of the field, though the p's now have different values.

Now suppose the field to change adiabatically. If the accelera-

* See Born, *Vorlesungen über Atommechanik*, pp. 238, 239.

Digitized by Google

tions ($\ddot{x}$, $\ddot{y}$, $\ddot{z}$) are the same as when the field is constant, we should have from (i) and (ii)

$$\dot{p}_x = -\frac{\partial H}{\partial x} - \frac{e}{c}\frac{\partial A_x}{\partial t}, \quad \dot{x} = \frac{\partial H}{\partial p_x} \qquad \text{.....(v)}$$

and the Hamiltonian equations would not be satisfied. Actually, the change in the magnetic field $\mathbf{H}$ must be accompanied by an electric field $\mathbf{E}$ satisfying

$$\operatorname{curl} \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{H}}{\partial t}, \qquad \text{.....(vi)}$$

the general solution of which is

$$\mathbf{E} = -\frac{1}{c}\frac{\partial A}{\partial t} - \operatorname{grad} \phi.$$

The scalar potential ϕ can have different values according to the way in which the change in the magnetic field is carried out. Owing to this electric field, $m\ddot{x}$ will not be the same function of the (x, y, z)'s and $(\dot{x}, \dot{y}, \dot{z})$'s as when the magnetic field is constant, but will be increased by $-e\mathbf{E}_x$. From (ii), $\dot{p}_x$ must be increased by the same amount. Hence, from (v)

$$\dot{p}_x = -\frac{\partial H}{\partial x} + e\frac{\partial \phi}{\partial x} = -\frac{\partial H'}{\partial x}, \quad \dot{x} = \frac{\partial H}{\partial p_x} = \frac{\partial H'}{\partial p_x},$$

where

$$H' = H - e\phi.$$

Hamilton's equations thus still hold with the new Hamiltonian H'.

Equation (iii) can no longer be exactly true, since $\mathbf{H}$ cannot be uniform on account of the equation

$$\frac{\partial^2 \mathbf{H}}{\partial x^2} + \frac{\partial^2 \mathbf{H}}{\partial y^2} + \frac{\partial^2 \mathbf{H}}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 \mathbf{H}}{\partial t^2} = 0.$$

This equation shows, however, that the variations in $\mathbf{H}$ are of the order T^{-2}, where T is the time taken to establish the field, so that the error introduced in equation (iv) is of the order T^{-2}. Also, on account of the relation

$$\operatorname{curl} \mathbf{H} = \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t},$$

$\dot{\mathbf{E}}$ is of the order T^{-2}, and $\mathbf{E}$ and ϕ are of the order T^{-1}. Hence H' differs from $H_0 + e/2mc.|\,\mathbf{H}\,|\,p_\phi$ only by terms of the order of magnitude T^{-1} or $\dot{\mathbf{H}}$, so that the ordinary proof of the adiabatic invariance of the quantum integrals can now be applied.

Effect of rapidly varying fields.

Sommerfeld* states that if the magnetic field is introduced suddenly, the momentary velocity of an electron will not be affected.

* *Atombau und Spektrallinien*, 4th edition, p. 403.

Digitized by Google

Equation (vi) shows, however, that a sudden change in the magnetic field must be accompanied by an impulsive electric field, which will in general cause a finite change in the velocity of the electron.

One way of establishing the field is by means of a plane wave advancing in the direction of the x axis with the velocity of light and having $\mathbf{E}_y = \mathbf{H}_z =$ a constant C behind the wave front, the other components of $\mathbf{E}$ and $\mathbf{H}$ vanishing. As this wave would establish a permanent electric field as well as a magnetic one, we must suppose another similar wave advancing in the opposite direction and having $\mathbf{E}_y = -\mathbf{H}_z = -C$ behind its wave front, so that the two waves together establish a uniform magnetic field $\mathbf{H}_z = 2C$ and no electric field. If the two wave fronts are timed to arrive at the origin (where the atomic nucleus is supposed to be situated) at the same time, an electron at (x, y, z) will be subject to the electric field $\mathbf{E}_y = -C$ for a time interval $2x/c$ (or to the electric field $\mathbf{E}_y = C$ for the time interval $-2x/c$ if x is negative), this being the difference between the times of arrival of the two wave fronts at (x, y, z). We are neglecting the displacement of the electron during the time $2x/c$, which is permissible if the velocity of the electron is small compared with that of light. The electron thus receives an impulse $2xeC/c = xe\mathbf{H}/c$ in the direction of the y axis. This can be resolved into a radial impulse $xye\mathbf{H}/rc$ and an impulse $x^2e\mathbf{H}/rc$ tangentially about the z axis, where $r^2 = x^2 + y^2$.

Now suppose that the field is established by such waves coming uniformly and simultaneously from all directions in the xy plane instead of only along the x axis. The radial impulse is now given by the average value of

$$(x \cos \alpha + y \sin \alpha)(y \cos \alpha - x \sin \alpha)\, e\mathbf{H}/rc$$

which is zero, and the tangential impulse by the average value of

$$(x \cos \alpha + y \sin \alpha)^2\, e\mathbf{H}/rc$$

which is $\frac{1}{2}r^2e\mathbf{H}/rc$. The angular velocity imparted to the electron is thus $\frac{1}{2}e\mathbf{H}/mc$, which is just the Larmor precession. Hence the quantum integrals will remain invariant when the field is suddenly established in this way, since the motion is the same as when the field is established adiabatically.

Any method of establishing the field that is symmetrical about an axis through the nucleus can be resolved into waves of the type just considered, so that the quantum integrals must also be invariant in this more general case.

Digitized by Google

A photographic method of determining the mobility of recoil atoms.
By G. H. Briggs. (Communicated by Dr J. Chadwick.)

[*Read* 7 December 1925.]

(Plate I.)

The experiments described in this paper were carried out in the University of Sydney during November 1924 and, owing to the short time which was available, are of a purely preliminary character. The following account is given as they are not likely to be resumed in the immediate future.

The method employed is in principle a modification of Zeleny's, using parallel plates, an arrangement which has been used by Nolan*, Erikson† and Tyndall and Grindley‡. In the experiments to be described instead of using an ionisation method advantage is taken of the photographic action of the recoil particles to determine their distribution. The experiments were begun in order to see if any difference in the mobilities of Thorium X, Thorium A and Thorium B could be detected.

The main part of the apparatus was a box built from two brass plates 20 cm. long and 7·4 cm. wide, separated at the longer edges by ebonite so that the internal cross section was 5·4 cm. × 1·0 cm. Air from a large gasometer after passing over pumice moistened with sulphuric acid and through a bottle of about 16 litres capacity filled with calcium chloride was led into the box through a tube 6 cm. in diameter and 80 cm. long. The section of this tube was gradually changed to fit the rectangular opening of the box, care being taken to smooth out irregularities. The other end of the box was open to the atmosphere. The air used in the experiments was certainly not thoroughly dried and its degree of dryness varied in different experiments.

About half-way along the box, a strip of platinum foil 3 mm. wide and 3·5 cm. long was soldered to the bottom brass plate with its length perpendicular to the direction of the air current. It was desired to obtain a narrow line of Mesothorium bromide along the centre of this strip. The method finally adopted was to rule with a drawing pen a line of the bromide, dissolved in a small quantity of water, on the foil which had previously been slightly platinized to reduce the effect of surface tension. The actual width was about 0·6 mm. and the amount of Mesothorium deposited was equivalent to about 0·08 mg. of Radium. The writer§ has recently shown

* Nolan, *Proc. Roy. Irish Acad.* vol. 35, p. 38. (1920.)
† Erikson, *Phys. Rev.* vol. 17, p. 400. (1921.)
‡ Tyndall and Grindley, *Phil. Mag.* vol. 47, p. 689. (1924.)
§ Briggs, *Phil. Mag.* vol. 50, p. 600. (1925.)

Digitized by Google

that such a source gives Th. X, Th. A, and Th. B, with an efficiency of recoil of about 3 to 4 per cent. With suitable electric fields the recoil atoms travelled across and were caught on a circular disc which fitted flush into the upper plate. With no air current a line of activity immediately above the source was formed, while with a steady flow the line was displaced down stream. If the three recoiling substances have different mobilities the displaced line would be expected to have three components. The distribution of the activity collected on the disc was found from the impression produced when the disc was placed face down on a photographic plate. The electric fields employed in various experiments ranged from 75 to 140 volts per cm. and the maximum displacement of the recoil atoms was about 2·5 cm. Exposures of 30 minutes without, and one hour with the air current, gave enough activity to produce a strong photographic impression when the activated disc was left on a photographic plate for some hours.

The type of photographic impression actually obtained is shown in Plate I. In the original photographs the displaced line consists chiefly of a number of elongated spots with their long axis parallel to the direction of the air stream, and it is seen that they correspond to dark spots in the undisplaced line. They are simply due to points on the source from which recoil is greater than from neighbouring regions. The undisplaced line is only about 0·5 mm. wider than the source so that diffusion and self repulsion are practically negligible. This is also clear from a consideration of the width of the elongated spots in the deflected line. The edge of the displaced band is much sharper on the high mobility side than on the other.

The relative width of the displaced band to the total displacement was found to be constant in all the experiments. These extended over a range of air currents varying by a factor of three. This is further evidence that the width of the band is not due to eddies. Again, eddies would make the broadening symmetrical about the mean deflection.

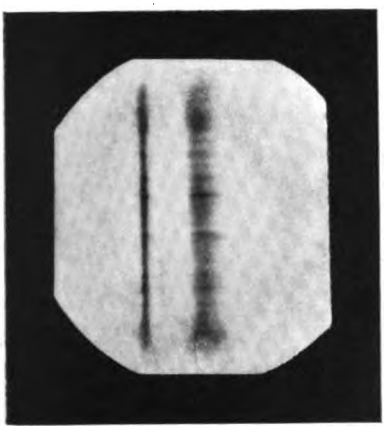
For plates so wide that the effect of the sides of the box is negligible the mobility is given by

$$k = \frac{q.d^2}{x.a.v},$$

where x is the displacement, v the voltage, d the separation of the plates, q the volume of air flowing per second and a the area of cross section. In the present experiments the photographs show that the displacement was constant over at least 3·5 cm. (the length of the source) of the total 5·4 cm., thus indicating that the effect of the sides of the box was small.

An attempt was made to see if the distribution of the three

Digitized by Google



a *b* Full size

a, undisplaced line

b, displaced line

Digitized by Google

Digitized by Google

substances Th. X, Th. A, Th. B in the band of activity was the same for each substance by exposing a freshly activated disc to photographic plates for a short time immediately after the experiment, then again after some hours and finally after several days. No difference in the photographic impression was observed, but it is not felt that the evidence on this point is at all conclusive.

Fig. 1 shows the blackening of one of the photographs measured in the Cavendish Laboratory with a microphotometer. Kinoshita* found for two kinds of photographic plates that the blackening produced by α-rays follows a law of the form $D = D_{\max.} - e^{-CN}$, N being the number of α particles and C a constant. The writer has found that a law of the same form holds for Ilford Process

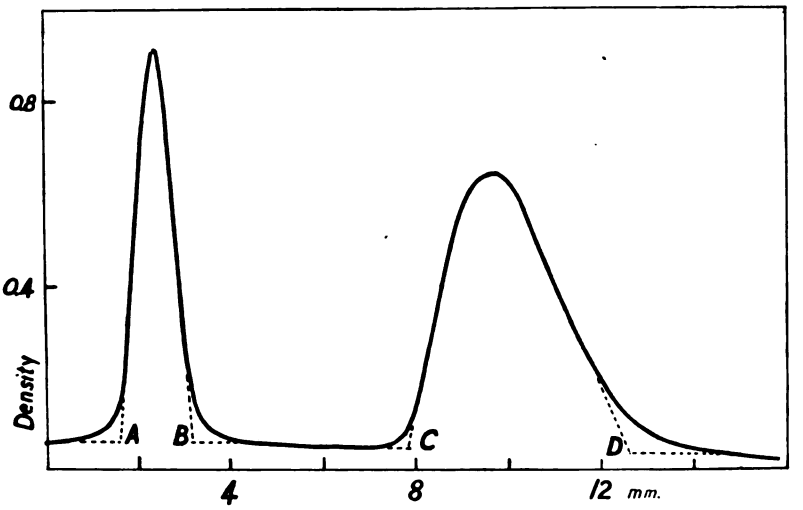


Fig. 1

plates and that up to a density of unity it is practically a linear relationship. Although the plates used in the present work, Imperial Eclipse, have not been calibrated there is little doubt that the density curve, Fig. 1, gives the distribution of α-ray activity quite accurately.

The values of the maximum and minimum mobilities have been calculated from the distances AC and BD (Fig. 1) on the curves found with the microphotometer. These are given in Table I and also the mobility corresponding to the peak of the curve.

In these experiments the ionisation from the source is so intense that with the strongest fields used it is only about 80 per cent. saturated, so that these electric fields may be disturbed by amounts

* Kinoshita, *Proc. Roy. Soc.* vol. 83, p. 432. (1910.)

Digitized by Google

sufficient to cause errors of a few per cent. The results are given without any correction. By using a point source of recoil atoms the amount of this ionisation could probably be reduced to less than one hundredth of its present value.

Erikson has recently measured the mobility of the active deposits of Actinium*, Thorium† and Radium†, using a method which is in principle similar to the present one, but detecting the active deposit by an ionisation method. He obtains a curve with two peaks, one corresponding to a mobility of 4·35 and the other to 1·55. Erikson also notes that the width of the curve for the group of slower particles is greater than is to be expected from the turbulence of the air and the width of the stream of particles

Table I.

Volts	Mobility		
	Max.	Peak	Min.
118	1·91	1·57	1·29 cm./sec./volt/cm.
127	1·87	1·61	1·30
134	1·80	1·53	1·18
77	1·91	1·67	1·26
186	1·80	1·66	1·32
140	1·77	1·49	1·10
Mean	1·84	1·59	1·24 ± 0·05

entering his apparatus. Calculating approximately the maximum and minimum mobility from Erikson's curve for the slower particles without attempting to correct for spread due to turbulence, etc. one gets a range appreciably greater than found in the present experiments but with practically the same peak value. It appears, therefore, that the writer's results are in general agreement with Erikson's in regard to the low mobility group of recoil particles. The fact that the present experiments do not bring to light the particles of mobility 4·35 cm. per sec. needs some consideration.

In Erikson's experiments with Radium and Thorium active deposit the age of most of the recoil particles before entering the box was many seconds or even minutes, while in his experiments with actinium emanation their age was probably between 0·01 and 0·002 sec. whereas in the present experiments they are produced actually in the field and cross the box in about 0·01 sec. Age then does not appear to be the determining factor. Among the different conditions in the two experiments those which seem most likely to be of importance are the degree of ionisation in the exposure apparatus and the intensity of the electric fields; for the

* Erikson, *Phys. Rev.* vol. 24, p. 622. (1924.)
† Erikson, *Phys. Rev.* vol. 26, p. 629. (1925.)

Digitized by Google

fields employed by Erikson, of the order of 1000 volts per cm., were roughly 6 or 8 times greater than those in the present experiments and the ionisation in his exposure apparatus was much smaller. This suggests then that the recoil ion with a mobility of 4·35 cm. per sec. is able to exist only in strong electric fields or in the absence of strong ionisation, though it should be noted here that in Franck's* experiments which gave the mobility as 1·56 cm. per sec. fields of about 600 volts per cm. were used.

The sharpness of the least deflected edge of the band found in the present experiments strongly suggests that the particles set out with a certain mobility and after going a greater or less distance may change into particles of smaller mobility. Those that fail to change arrive at the inner edge and those that change at the beginning reach the plate at the outer edge of the band. On this view the mobilities calculated from the position of the edges of the band are of more significance than that for the peak.

* Franck, *Verh. Deutsch. Phys. Gesell.* vol. 7, p. 397. (1909.)

Digitized by Google

The quenching of resonance radiation and the breadths of absorption-lines. By R. W. DITCHBURN, B.A., Trinity College.

[*Read* 7 December 1925.]

In a paper by Slater* a new theory of radiation is discussed and the following explanation of the breadths of spectral lines is given: "An atom in the ith state has a probability P_i of suffering in unit time a transition. Thus there is a probability P_i that the vibrations of each of the oscillators will simultaneously cease. But we shall assume that, in addition to this probability P_i of ceasing its oscillation altogether, each oscillator has also an independent probability P_j of suffering an interruption in which it ceases its oscillation as if it were leaving the state, but immediately begins again as if it were entering the same state" (i.e. with an arbitrary phase difference). "*This term P_j is the same as the probability that an atom in the jth state will leave that state.*" The total probability of interruption of vibration is thus $(P_i + P_j)$ which is symmetrical with regard to the two end states. This makes the breadths of absorption and emission lines equal and so satisfies Kirchhoff's law.

In the case, however, where j is the normal state of the atom P_j is vanishingly small (except when the atom is in a radiation field of enormous density—or subject to similarly strong electron bombardment). We should therefore expect that in this case, the transition probability would be the same as the total probability of interruption of oscillation; i.e. that practically every disturbance which interrupted a coherent train of waves would also cause a transition.

The experiments of Fuchtbauer, Joos and Dinkelacker† on the broadening of the 2536 line of Hg by the addition of various foreign gases enable us to calculate the probability of interruption of coherent-wave trains. The experiments of Stuart‡ on the quenching of the 2536 resonance radiation by foreign gases enable us to calculate the transition probability under the conditions of Fuchtbauer's experiments. A comparison of the two sets of results forms a test of the theory.

(A) *The Quenching of Resonance Radiation.*

The rate of resonance radiation is given by Slater§ as

$$N\left[\frac{A_{ij}B_{ij}(\nu)\rho(\nu)h\nu}{A_{ij}+2B_{ij}(\nu)\rho(\nu)+K_{ij}}\right],$$

* *Phys. Rev.* vol. 25, p. 395.
† *Ann. d. Phys.* vol. 71, p. 204.
‡ *Zs. f. P.* vol. 32, p. 262.
§ Equation 20, p. 419, *loc. cit.*

Digitized by Google

where

N is the total number of atoms of the radiating substance, and

A_{ij} is the probability of spontaneous transition (in unit time) from state i to state j.

K_{ij} is the probability of transition from state i to state j by collisions of the second type.

$B_{ij}(\nu)\,\rho(\nu)$ is the probability of transition from state i to state j stimulated by the radiation field whose density is $\rho(\nu)$.

Thus K_{ij} is proportional to the pressure of the foreign gas present but A_{ij} and B_{ij} are independent of this pressure. If we define I_p as the ratio of the intensity of resonance radiation when the pressure of the foreign gas is p to the intensity of the resonance radiation when it is zero, we have

$$I_p = \frac{A_{ij} + 2B_{ij}(\nu)\,\rho(\nu)}{A_{ij} + 2B_{ij}(\nu)\,\rho(\nu) + (K_{ij})_p} \qquad \text{.....(1),}$$

where $(K_{ij})_p$ is the value of K_{ij} when pressure of foreign gas is p. "For all ordinary cases $B_{ij}(\nu)\,\rho(\nu)$ is negligible in comparison with A_{ij}," so that we may write

$$I_p = \frac{A_{ij}}{A_{ij} + (K_{ij})_p} \qquad \text{.....(2).}$$

It is obvious that if I_{p_0} is $\frac{1}{2}$ then $(K_{ij})_{p_0}$ is equal to A_{ij}, p_0 being the pressure required to reduce the resonance radiation to half-value. We can calculate $(K_{ij})_{p_0}$ on the assumption that every kinetic-theory collision is effective in causing a transition from state i to state j. The reciprocal of $(K_{ij})_{p_0}$ is the mean time between two collisions and the reciprocal of A_{ij} is the mean life of the excited atom. For certain gases Stuart finds that, taking the ordinary values of the atomic radii, the mean time between two kinetic theory collisions at the half-value pressure is greater than the mean life of the excited atom (as determined by Wood and Ellet, Hanle, Wien and Tolman). This he explains on the assumption that the radius of the excited mercury atom is much greater than that of the normal atom. There is no direct method of determining exactly what the radius of the excited atom is, so Stuart assumes that in the case of oxygen every collision with the excited atom is a collision of the second type, i.e. is effective in causing a transition. This gives the radius of the excited atom as 3·4 times that of the normal mercury atom and he then calculates the percentage of collisions (within this radius) which are collisions of the second type, in the cases of other gases. His results are given in the following table.

Digitized by Google

TABLE I.

	%
Oxygen (standard)	100·0
Carbon monoxide	80·0
Hydrogen	70·0
Carbon dioxide	20·0
Water vapour	10·0
Nitrogen	1·3
Argon	0·2
Helium	0·03

It is obvious that the above are only relative "efficiencies of collision." If the radius of the excited mercury atom is greater than it has been assumed to be, then the efficiency for oxygen will be less than 100 per cent. and all the other percentages will have to be reduced accordingly.

(B) *The Breadths of Spectral Lines.*

The half-value width of the absorption lines is given by Slater* as equal to

$$\frac{1}{\pi}(P_i + P_j).$$

Making the same assumptions as before this reduces to

$$\frac{1}{\pi}[A_{ij} + (K_{ij})_p] \qquad \ldots\ldots(3).$$

When p is of the order 1 atmosphere or greater $(K_{ij})_p$ is fairly large compared with A_{ij} and (assuming that $(K_{ij})_p$ is proportional to p) the breadth should then be proportional to the pressure. Fuchtbauer's experiments were made at pressures from 10 to 50 atmospheres and he found that the breadth is in fact proportional to the pressure.

Since Stuart's results give the number of collisions of the second kind for the different gases and since according to the theory this is also the number of collisions which are effective in interrupting coherent wave-trains, we can now calculate (from equation (3)) the coefficient of this proportionality. The comparison between the observed values and those calculated on the above assumption is given in the first two columns of Table II†. The third column gives the values calculated on the assumption that every collision (within the radius of the excited molecule found by Stuart) is effective in interrupting coherent wave-trains. The coefficient is given as the width for one atmosphere pressure. The widths are expressed in frequency measure and the unit is sec^{-1}.

* p. 416, *loc. cit.*

† Column (i) gives the observed values.

Digitized by Google

TABLE II.

	(i)	(ii)	(iii)
Oxygen	$7{\cdot}9 \times 10^9$	$7{\cdot}0 \times 10^9$	$7{\cdot}0 \times 10^9$
Hydrogen	$12{\cdot}3 \times 10^9$	$10{\cdot}5 \times 10^9$	$15{\cdot}0 \times 10^9$
Carbon dioxide	$13{\cdot}1 \times 10^9$	$1{\cdot}2 \times 10^9$	$6{\cdot}0 \times 10^9$
Water vapour	$10{\cdot}6 \times 10^9$	$0{\cdot}6 \times 10^9$	$6{\cdot}0 \times 10^9$
Nitrogen	$8{\cdot}3 \times 10^9$	$0{\cdot}078 \times 10^9$	$6{\cdot}0 \times 10^9$
Argon	$9{\cdot}6 \times 10^9$	$0{\cdot}01 \times 10^9$	$5{\cdot}1 \times 10^9$

The disagreement between columns (i) and (ii) shows the discrepancy between theory and experiment. This discrepancy is, perhaps, seen more clearly if the results are treated in the following way. Comparing columns (i) and (iii) of Table II we see that (except in the case of hydrogen) the number of collisions which interrupt coherent wave-trains is apparently greater than the number of kinetic theory collisions. This means that the effective radius for this purpose is greater than that found by Stuart. On the assumption that every collision with a carbon-dioxide molecule is effective in ending a train of coherent vibrations we obtain a radius for the mercury atom of $9{\cdot}6 \times 10^{-8}$ cm., which is 5·5 times the kinetic theory radius for the normal atom. Taking this value we may now calculate (*a*) from Stuart's results, the percentage of collisions within this radius which are effective in causing transitions; (*b*) from Fuchtbauer's results, the percentage of collisions within this radius which are effective in ending a train of coherent waves. The results for the different gases are given in the following table:

TABLE III.

	(*a*)	(*b*)
	%	%
Hydrogen	23·0	24
Water vapour	3·3	57
Nitrogen	0·51	52
Oxygen	47·0	53
Argon	0·08	73
Carbon dioxide	9·2	100
Helium	0·009	(No observations)
Carbon monoxide	37·0	—

Thus we see that in the case of carbon dioxide only about 10 per cent. of those collisions which are effective in ending a train of coherent waves are also effective in causing transitions and this is typical of the other gases.

This seems quite incompatible not only with the theory

Digitized by Google

referred to above but also with any form of the virtual-oscillator theory which satisfies Kirchhoff's law by assuming that the probability of a collision which ends a coherent wave-train and does not cause a transition is equal to the probability of an atom leaving the lower state. This conclusion would not be essentially altered even if one makes the somewhat unlikely assumption that only part of the breadth of an absorption-line is due to the molecular collisions. The reasoning would apply equally to any other cause of interruption of vibration.

Kirchhoff's law can be satisfied by making the total probability of ending a train of coherent waves equal to any symmetrical function of P_i and P_j. In addition to satisfying Kirchhoff's law this function would have to give the statistical conservation of energy and momentum and also to explain the above experiments quantitatively. The simplest assumption is that the total probability of ending a coherent wave-train is $(P_i + P_j + P)$, where P is a constant for all states of the atom. By adjusting the value of P for various gases Fuchtbauer's and Stuart's results could be explained. This probability P would have to be a collision probability which would be the same for an atom in the normal state as for one with an electron in an orbit of very large radius. It is possible to give a physical meaning to such a probability by assuming that in order that vibrations may be interrupted it is necessary for an atom to pass within a certain distance of the particular electron which is vibrating or, alternatively, within a given distance of the nucleus of the vibrating atom. The writer has not investigated the energy and momentum relations.

The difficulty revealed above is not met with if one approaches these results from the point of view of the older light quantum theory. There is one type of collision which interrupts the atomic oscillation and another which causes transitions. There is no reason why these should not possess totally different probabilities (as in fact they do). Kirchhoff's law is satisfied as long as the total chance of interruption of vibrations is the same for an emitting as for an absorbing atom. This may be secured, as far as interruption by collision is concerned, by making the effective collision radius equal to some kind of mean between the radius of the normal atom and that of the excited atom. The details of this explanation are, of course, obscure but that is because the light-quantum theory is unable to give any detailed picture of the emission and absorption process.

It is interesting to consider the physical difference between these two types of collision. There is, as Stuart observes, no obvious relation between the efficiency of collisions of the second kind (i.e. transition collisions) and the atomic field, polarity or any other property of the colliding atom. It seems probable that this is

Digitized by Google

a very intimate type of collision and possibly a temporary compound is formed. The series-electron may go into some kind of linking orbit and the probability of this happening would naturally depend on the fitting together of the electronic systems of the two atoms. The temporary compound would soon disintegrate (either spontaneously or under the influence of other atoms) and the energy of excitation would thus go into thermal energy. It is natural to suppose that in some cases the compound is permanent and it would be interesting to test for a (comparatively) rapid interaction of oxygen and mercury (at ordinary temperatures) under the influence of the resonance light.

The efficiency of the interruption of vibration collisions seems to depend on the molecular weight of the foreign gas. A reference to Table III above will show that this efficiency increases gradually with the molecular weight except in the case of water-vapour.

The experimental difficulties of determining the breadths of the absorption lines are greater in the case of water than for the other substances because it is a vapour and the others are "permanent" gases. Moreover, the values obtained (by different methods) for the molecular radius of the water-vapour molecule differ very much and if one uses a different set of values of the molecular radii the efficiency for water becomes a little less than that for nitrogen, the other efficiencies remaining in the same order as before. Thus it is fairly probable that the molecular weight is one factor, at least, in determining the efficiency of this type of collision.

The high efficiency of the heavy molecule may be due to its high average momentum. Thus a heavy molecule should probably create a greater disturbance in the electronic system of the molecule with which it collides than a light one would.

On the other hand, it may be that it is necessary for the two atoms to remain within a given radius of each other for a certain minimum time in order that the orbital motion of the series electron may bring it into such a position that its motion is affected by the other atom. The heaviest atoms having the lowest mean-square velocities would thus be the most efficient.

Thus an excited atom can collide with a normal atom in at least two distinct ways: (i) in which the phase of its electronic vibration is changed; (ii) in which the energy of its excitation is given up to the colliding atom. It would seem at first sight that one might have another type of collision in which the excited atom would be stimulated to give up its energy, not to the other atom, but to the aether. This idea of stimulated emission is attractive in many ways but its existence seems to be ruled out by the following considerations. Suppose a mixture of a very efficient and a very inefficient gas (say oxygen and helium) to be used in the experiment of Stuart. We should then expect the mixture to be less effective than the

Digitized by Google

oxygen by itself (at the same partial pressure) because the helium molecules would collide with some of the excited atoms and cause them to give up their energy as radiation before a collision with an oxygen atom could turn it into heat energy. This effect should be quite appreciable in the case of air (considered as a mixture of oxygen and nitrogen). Stuart finds, however, that the efficiency of the mixture is equal (within experimental error) to that to be expected on an additive law. This shows that as regards mercury and nitrogen, at any rate, this stimulated emission does not play any important part. We may thus conclude that a collision can shorten the life of an excited atom only by turning its energy of excitation into heat energy and not by transforming it into resonance radiation.

I am indebted to Mr R. H. Fowler for opportunities of discussing this work with him.

Digitized by Google

Notes on the electrical counter. By K. G. EMELÉUS, B.A. St John's College. (Communicated by Mr E. V. APPLETON.)

[*Read* 7 December 1925.]

In a recent communication on the subject of the electrical counter, it was suggested that when an ionizing particle enters, it brings about a momentary brush discharge from the smaller of the two electrodes, and that the transitory nature of this discharge is due to a local increase of pressure in the gas*. Some further experiments concerning its use and mode of action have now to be recorded. In most of these, the point counter devised by Geiger† was employed. Its outer cylinder was maintained at a high potential by a motor-generator or battery, and the needle was earthed, usually through a high resistance. In general, the point was negative with respect to the case, and the counter was filled with air at atmospheric pressure, in contact with phosphorus pentoxide.

1. *Examination of the point during counting.*

A counter was made with which the point (*P*, Fig. 1) could be watched with a microscope‡, whose objective was just outside a mica window (*M*), fixed with wax over a hole in the case. It could be exhausted through a side-tube (*T*). α-particles entered the

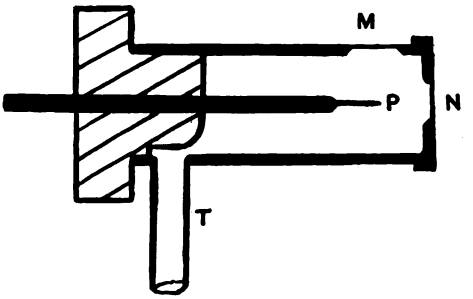


Fig. 1

chamber through a second mica window (*N*), and were recorded in the usual way by an electrometer. It was seen that when the case was positive, each particle gave rise to a flash at the extreme end of the needle. This had the appearance of a typical negative brush discharge; a faintly luminous cone of about 30° spread out

* Emeléus, *Proc. Camb. Phil. Soc.* vol. 22 (1925), p. 676.
† Geiger, *Verh. d. D. Phys. Ges.* vol. 15 (1913), p. 534.
‡ Cf. also Wulf, *Phys. Zeit.* vol. 26 (1925), p. 382.

Digitized by Google

axially from the tip of the needle into the body of the gas, and became brighter and longer when the voltage was raised or the pressure lowered. There was also present a luminous collar round the metal, at about 0·1 mm. back from the tip at atmospheric pressure, and similar to the collar often observed with a sustained discharge at pressures of the order of 1 mm. Hg. If the potential was raised above its upper limit for efficient counting, the natural discharges from the point were seen to be similar to those already described. The effects in air and hydrogen were similar, but at any one pressure a smaller potential was needed for the latter gas, and the visible discharge occupied a smaller volume. Identical flashes were produced by β-particles.

The motor-generator was not sufficiently steady to keep the counter working over the very limited range for which it would respond to an ionizing particle when the point was at a higher potential than the case, but the natural discharges were typical positive brushes from the point*.

2. *Influence of a magnetic field.*

A counter which worked satisfactorily between 1400 volts and 1700 volts when its point was negative, was sheathed in thin ebonite, and put between the poles of an electromagnet. Fields of up to 4000 gauss were applied both longitudinally and transversely. It was found that the magnetic field affected neither the magnitude of the throw of the electrometer, nor the number of particles recorded, using α-rays from polonium, and γ-rays from radium bromide and its products: the latter would produce secondary β-particles inside the counting chamber. This absence of effect is consistent with the view that the main action takes place in the immediate neighbourhood of the point, where the force on an ion is already that due to an electric field of the order of 10^5 volts/cm.† It is of course possible to prevent an ionizing particle from ever entering the counter, by applying a magnetic field transverse to its path‡.

3. *The circuit resistance.*

It has been usual to employ an enclosed mixture of alcohol and xylol, with platinum leads, as a high-resistance leak between the counter and earth. This is not readily made in a variable form, and a thermionic valve has been found to be a satisfactory substitute. The plate and grid are connected to the needle of the counter, and one end of the filament is earthed. Its effective resistance is controlled by the temperature of the filament.

* Cf. Wulf, *loc. cit.*
† Emeléus, *Proc. Camb. Phil. Soc.* vol. 22 (1925), p. 676.
‡ Emeléus, *loc. cit.* vol. 22 (1924), p. 400.

Digitized by Google

With the usual circuit, the obvious function of this resistance is to make it possible to record the particles with an electrostatic instrument. We have pointed out, however, that under some circumstances it must also be present to extinguish the discharge*, and on one occasion when it was short-circuited, a 1 amp. fuse was blown. On the other hand, Kutzner has stated that it is not required to prevent arcing†, and in a circuit described by Wulf‡ the particles are indicated by clicks in a pair of telephones in series with the counter, the total resistance being only some 2000 ohms.

The essential difference is that in our experiments the gas in the counter was at a lower pressure than in those of Kutzner and of Wulf. The measurements of the fall of potential due to a single particle show that at high pressures extinction of the discharge can be accounted for by a local increase of pressure in the gas at the point, whilst at low pressures it is the fall in the potential across the counter which brings about such a diminution of the field that ionization by collision stops. The local increase of pressure does not depend on the circuit constants, and in consequence the discharge is transient with or without a large charging resistance: the potential between the electrodes is, on the contrary, a function of the resistance, and if the latter is less than some critical value, the stable state of the system is one in which a relatively large steady current is flowing.

In order to be sure that in these experiments the difference was not due to the use of different counters, one was set up in which the pressure could be varied; the particles were recorded either by telephones in series, and by microscopic examination of the point, or by the electrometer method. As was expected, no large resistance was needed when the pressure was above 3 cm. Hg, particles being recorded in the absence of a resistance by flashes or clicks, and in presence of a resistance by flashes or by the electrometer. Below this pressure, a discharge, when once started by an ionizing particle or apparently spontaneously, proceeded in absence of resistance without further ionization of external origin. The resistance was now increased by having in series a carefully insulated valve and gradually reducing its filament current, and the system was found to pass through a state in which the continuous current was replaced by bursts of discharges to a final stage when no current flowed except in the normal course of counting.

The necessity for resistance at low pressures can be shown readily with a commercial discharge-tube filled with neon. If an "Osglim" lamp (*O*, Fig. 2) be set up in parallel with a condenser (*C*) and in series with a diode (*D*)—to act as a variable resistance—

* Appleton, Emeléus and Barnett, *loc. cit.* p. 448.
† Kutzner, *Zeit. f. Phys.* vol. 21 (1924), p. 281.
‡ Wulf, *loc. cit.*

Digitized by Google

then when suitably adjusted it will flash intermittently on exposure to γ-rays*. Both the capacity and effective resistance must be fairly high (e.g. 0·1 microfarad, and the filament of a D.E.R. valve just visible warm), and the voltage of the battery (B) has usually to be less than a volt below the potential difference at which regular capacity-resistance oscillations start. If now when the lamp is responding to β-particles in much the same way as a point counter, the filament current of the valve is gradually increased (or what is equivalent, the capacity C is reduced), the discharge through the neon proceeds from irregular flashing through a stage where it gives bursts of flashes ("spurious discharges"), till finally a discharge, when once started by exposure to γ-radiation, passes apparently continuously and without further external stimulus.

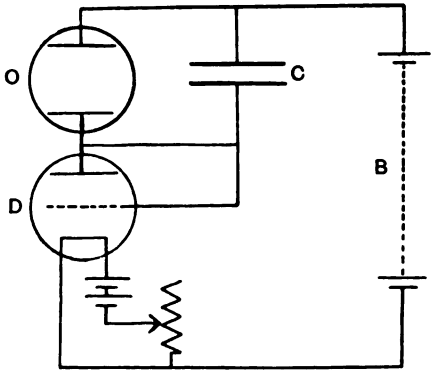


Fig. 2

4. *Reaction on the electrodes.*

Each particle, as it is recorded, gives a change in potential across the counter which is often of the order of 100 volts. The resulting change in the force of attraction between the two electrodes can be shown with the apparatus illustrated in Fig. 3. The counting point was the sharpened tip of a large steel needle (A), supported in an ebonite stand (B). The outer cylinder (C) was of thin aluminium foil, and was suspended from a glass frame by four silk threads (S), like the pan of a ballistic balance. The one end, opposite to an ebonite stop (D), was open, and the other end was covered with silver leaf. Leads from the motor-generator, and to the electrometer and earth, were brought up in glass tubes through an ebonite disc (E) and an earthed brass plate (F). To keep dust from the needle, which was negative, the apparatus was covered by a glass bell-jar.

As the potential of the cylinder was raised, it was displaced

* Baeyer and Kutzner, *Zeit. f. Phys.* vol. 21 (1924), p. 46.

Digitized by Google

longitudinally towards the base of the needle. γ-rays were registered between about 1300 volts and 1700 volts, and with each discharge, the cylinder was seen to jump back a short distance towards its undisplaced position. The size of the kick was very approximately what would be expected from the fall in potential indicated by the electrometer. For example, a potential difference of 1400 volts

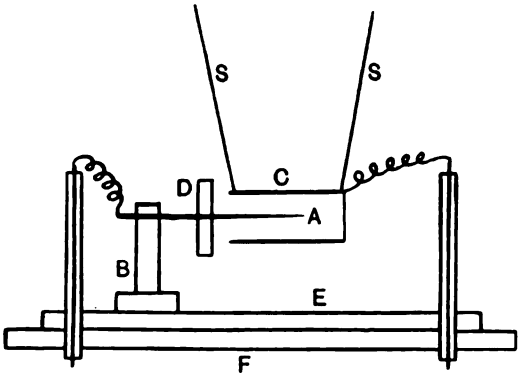


Fig. 3

gave a steady deflection of 0·8 mm. to the cylinder. The average ballistic displacement as a particle was recorded was 0·16 mm., corresponding to a fall in potential of a little over 100 volts: the electrometer reading was 60 volts. Similar relative motion was observed when the needle was suspended and the cylinder fixed, and when both could move.

It should not be difficult to use this effect to ring a bell by a relay. The working of a counter is, however, readily demonstrated to an audience by connecting a loud-speaker to the needle, through a valve amplifier.

5. *Hysteresis in the discharge.*

An attempt was made to show directly the existence of hysteresis in the characteristic curve of the counter*. To measure the current, the insulated pair of quadrants of a Dolezalek electrometer, used heterostatically, with a maximum sensitivity of 3000 mm./volt, was connected to the needle of the counter. The usual earthed shields were employed. To minimize insulation leaks in the counter itself, a cylindrical earthed guard-ring was inserted in the ebonite stopper, between the two electrodes. So long as the resistance between the needle and earth remained constant, the deflection of the electrometer was proportional to the current in the circuit.

* Appleton, Emeléus, and Barnett, *loc. cit.*

Digitized by Google

When a xylol-alcohol resistance was used, no detectable current passed unless a discharge had been excited, at at least 1400 volts, by an ionizing particle. On then decreasing the voltage, the current also diminished, Ohm's law holding approximately if the changes were made sufficiently rapidly. The current also decreased if the applied potential were kept unchanged, the fall being roughly exponential, with a half-period of about 5 minutes. The maximum steady current was 10^{-11} amp., the bursts accompanying counting being much larger, and of the order of 10^{-5} amp.* †

When a valve was used instead of the liquid resistance, it was less easy to obtain steady deflections of the electrometer, presumably because of small variations in the electron emission from the filament, but no hysteresis was found. The currents were of the same order as those which passed with the xylol after a high potential had been applied, but they were now determined uniquely by the voltage across the counter.

These results show that the hysteresis necessary for extinction of the discharge under these conditions is dynamical. The currents seem to be insulation leaks, or small continuous discharges from one or other of the electrodes, since control measurements showed that they were not due to induction from the high-tension leads or lighting mains, and were not due to radioactive preparations in the vicinity. Moreover, it appears that a xylol resistance, as here used, is liable to become polarized‡.

Since the currents were about the same whether the point was positive or negative, no support is given to the suggestion that the counting action is brought about by the release, by ions of opposite sign, of a space-charge held in the intense local field at the point. Some attempts to obtain such an action, with a commercial diode containing helium, and with a diode containing air, whose pressure was varied from 1/50th mm. Hg to atmospheric pressure, were without result.

It has been pointed out that when a condenser discharges periodically, recharging through a high resistance, some inertia is required to stabilize the oscillations. It was at first suggested that this was furnished by the natural inductance of the circuit*. It may, however, particularly at high pressures, be of mechanical origin, rather than electromagnetic, and due to movement of gas at the spark-gap.

6. *An effect of intense ionization.*

It has been noticed that when a strong radioactive preparation is brought near a counter at atmospheric pressure, with a positive

* Appleton, Emeléus and Barnett, *loc. cit.*
† Cf. also Geiger, *Zeit. f. Phys.* vol. 27 (1924), p. 7.
‡ Cf. Campbell, *Phil. Mag.* vol. 23 (1912), p. 668.

Digitized by Google

point, the number of particles recorded diminishes, whilst a continuous current passes*. This effect can also be produced at a pressure of a few cm. Hg, in which case the continuous current itself has a maximum value. Without at present attempting to explain it in detail, it may be pointed out that it is probably accompanied by electrical convection of the gas†, and that somewhat similar results follow if electrons are supplied to the cathode in an ordinary low-pressure discharge-tube‡.

Summary.

1. The brush discharges from a counting point have been examined with the microscope.

2. Magnetic fields have been shown not to affect counting under certain conditions.

3. The function of the series resistance has been further studied.

4. An approximate measurement has been made of the mechanical reaction on the electrodes.

5. An attempt has been made to detect hysteresis in the discharge, with negative results.

6. An effect of intense ionization has been noted.

The experiments described have been carried out at King's College, London. I have to thank Professor E. V. Appleton for his interest in their progress, and in particular for suggesting the use of a valve as a high resistance.

* Wulf, *loc. cit.*
† Trey, *Phys. Zeit.* vol. 22 (1921), p. 405.
‡ Wehnelt and Jachan, *Zeit. f. Phys.* vol. 31 (1925), p. 666.

Digitized by Google

On chains of two-two relations and the theory of elliptic functions. By Professor H. F. BAKER.

[*Received* 8 January, *read* 18 January, 1926.]

§ 1. The relation

$$a(1 + \phi^2\theta^2) + b\phi\theta + c(\phi^2 + \theta^2) = 0,$$

in which we suppose each of a, b, c, $a^2 - c^2$ to be different from zero, leads, from a value θ, to two values of ϕ, which we may denote by θ_1 and θ_{-1}. Each of these, put in place of θ, leads, beside the value θ of ϕ, to another value of ϕ, say, respectively, θ_2 and θ_{-2}. If θ_2 or θ_{-2} be put for θ, the same relation leads to two values of ϕ, say, θ_1 and θ_3 or θ_{-1} and θ_{-3}, respectively. And so on. It may happen that θ_n is the same as θ, in which case also θ_{-n} is the same as θ; this we may express by saying that the relation is closed, or that there is closure, after n links. It is the object of the present note to express in reduced form, in terms of a, b, c, the condition that this may be so. Evidently, if $n = pq$, the condition of closure after n links is satisfied when the condition for closure after p links (or also after q links) is satisfied. But there is a condition for closure after n links which is not satisfied for any less number of links; this is the condition which we call the *reduced* or *proper* condition for closure after n links, and it is this which we seek to express.

The matter has been dealt with by many writers; for this reason we limit ourselves as far as possible to a statement of results. We put $\beta = (a^2 - c^2)/b$, and express the reduced condition of closure by the vanishing of a rational integral homogeneous polynomial (with rational integer coefficients) in β, a, c, in the form $F_n(\beta, a, c) = 0$, for $n > 4$. For $n = 4$ the condition is $b(a^2 - c^2) = 0$, and for $n = 3$ it is $\beta + c = 0$. If, in terms of its prime factors, including, possibly, 2,

$$n = p^\alpha q^\beta r^\gamma \ldots,$$

we define $\psi(n)$ by

$$\psi(n) = n^2(1 - p^{-2})(1 - q^{-2}) \ldots,$$

the homogeneous polynomial $F_n(\beta, a, c)$ is, unless $n = 4$, of order

$$\tfrac{1}{8}\psi(n).$$

When n is odd this polynomial is irreducible; when n is even it is a product of three such rational polynomials. The values of $\frac{1}{8}\psi(n)$ for the following respective values of n are given by

$n = 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 24;$

$\frac{1}{8}\psi(n) = 1, \frac{3}{2}, 3, 3, 6, 6, 9, 9, 15, 12, 21, 18, 24, 24, 36, 27, 45, 36, 48.$

Digitized by Google

It can be proved that

$$n^2 - 1 = \sum_{d>1}^{n} \psi(d),$$

where d denotes in turn all the divisors of n, including n itself, which are greater than 1.

For each value of n we compute, beside the polynomial $F_n(\beta, a, c)$, two homogeneous polynomials, U_n and V_n; these are used in the computation of the polynomials $F_n(\beta, a, c)$ for greater values of n. When n is odd these are both of order $\frac{1}{4}(n^2 - 1)$, and both contain the term $\beta^{\frac{1}{2}(n^2-1)}$, the former with the sign $(-1)^{\frac{1}{2}(n-1)}$. This may be indicated by writing

$$U_n = (-1)^{\frac{1}{2}(n-1)}[\beta^{\frac{1}{2}(n^2-1)} + \ldots], \qquad V_n = \beta^{\frac{1}{2}(n^2-1)} + \ldots.$$

When n is even the similar statements are those denoted by

$$U_n = (-1)^{\frac{1}{2}n-1}[\beta^{\frac{1}{2}n^2} + \ldots], \qquad V_n = c[\tfrac{1}{2}n\beta^{\frac{1}{2}n^2-1} + \ldots],$$

both being of order $\frac{1}{4}n^2$. When the integer $\frac{1}{2}n$ is odd, U_n divides by β; when $\frac{1}{2}n$ is even, V_n divides by β.

§ 2. We now give a statement of results for $n = 3, 4, \ldots, 20$ and for $n = 24$. We use, for abbreviation,

$$P = \beta^4 - a^2c^2, \qquad Q = 2\beta^2 - a^2 - c^2,$$
$$M = P^4 - Q^4\beta^4a^2c^2, \qquad N = 2P^2 - Q^2(\beta^4 + a^2c^2),$$
$$H = 3P^4 - P^2Q^2(\beta^4 + a^2c^2) - Q^4\beta^4a^2c^2;$$

also $$\beta_1 = -P/Q, \qquad a_1 = ac, \qquad c_1 = \beta^2,$$

so that $$MQ^{-4} = \beta_1^4 - a_1^2c_1^2, \qquad NQ^{-2} = 2\beta_1^2 - a_1^2 - c_1^2,$$
$$HQ^{-4} = 3\beta_1^4 - \beta_1^2(a_1^2 + c_1^2) - a_1^2c_1^2.$$

It will be found that

$$P - a^2Q = (\beta^2 - a^2)^2, \qquad P - c^2Q = (\beta^2 - c^2)^2,$$
$$P - \beta^2Q = -(\beta^2 - a^2)(\beta^2 - c^2),$$

and $$P + \beta^2Q = 3\beta^4 - \beta^2(a^2 + c^2) - a^2c^2,$$

while

$$M + NP^2 = H, \qquad M - NP^2 = -(P^2 - Q^2\beta^4)(P^2 - Q^2a^2c^2).$$

In all cases $F_n(\beta, a, c)$ is an even function of a, or

$$F_n(\beta, a, c) = F_n(\beta, -a, c).$$

Also it may be noted that the terms containing the highest power of β in M, N, H are, respectively, β^{16}, $-2\beta^8$, $-\beta^{16}$.

$n = 1$ $\quad U_1 = 1, \qquad V_1 = 1.$

$n = 2$ $\quad U_2 = \beta, \qquad V_2 = c.$

Digitized by Google

$n = 3$ $F_3(\beta, a, c) = \beta + c, \quad U_3 = -(\beta^2 - a^2), \quad V_3 = \beta^2 - c^2.$

$n = 4$ $F_4(\beta, a, c) = \beta c, \quad U_4 = -P, \quad V_4 = Q\beta c.$

The $F_4(\beta, a, c)$ is purely formal.

$n = 5$ $F_5(\beta, a, c) = \beta^3 - \beta^2 c - \beta c^2 + a^2 c.$

$$U_5 = (\beta^2 - a^2) P - a^2 (\beta^2 - c^2) Q,$$
$$= \beta^2 (\beta^2 - a^2)^2 - a^2 (\beta^2 - c^2)^2,$$
$$V_5 = (\beta^2 - c^2) P - c^2 (\beta^2 - a^2) Q,$$
$$= \beta^2 (\beta^2 - c^2)^2 - c^2 (\beta^2 - a^2)^2.$$

$n = 6$ $F_6(\beta, a, c) = (\beta - c)(\beta^2 - a^2)$. And we have

$$F_3(\beta, a, c) F_6(\beta, a, c) = QF_3(\beta_1, a_1, c_1).$$

$$U_6 = \beta (P^2 - Q^2 a^2 c^2),$$
$$V_6 = c (\beta^2 - a^2)(\beta^2 - c^2)[3\beta^4 - \beta^2 (a^2 + c^2) - a^2 c^2].$$

$n = 7$ $F_7(\beta, a, c) = (\beta^2 - c^2) P + (\beta^2 - a^2) \beta c Q,$

$$= \beta^6 - \beta^4 c^2 - \beta^2 a^2 c^2 + a^2 c^4$$
$$+ \beta c [2\beta^4 - \beta^2 (3a^2 + c^2) + a^2 (a^2 + c^2)],$$
$$U_7 = -F_7(\beta, c, a) F_7(\beta, -c, -a),$$
$$V_7 = F_7(\beta, a, c) F_7(\beta, -a, -c).$$

$n = 8$ $F_8(\beta, a, c) = PQ,$

$$U_8 = -(P^4 - Q^4 \beta^4 a^2 c^2),$$
$$V_8 = \beta c PQ [(\beta^4 + a^2 c^2) Q^2 - 2P^2].$$

$n = 9$ $F_9(\beta, a, c) = \beta (P^2 - a^2 c^2 Q^2) - c (P - a^2 Q)(P + \beta^2 Q),$

$$U_9 = (\beta^2 - a^2)[\beta^2 (P^2 - a^2 c^2 Q^2)^2$$
$$- a^2 (P - c^2 Q)^2 (P + \beta^2 Q)^2],$$
$$V_9 = (\beta^2 - c^2)[\beta^2 (P^2 - a^2 c^2 Q^2)^2$$
$$- c^2 (P - a^2 Q)^2 (P + \beta^2 Q)^2].$$

$n = 10$ $F_{10}(\beta, a, c) = F_5(\beta, -a, -c) F_5(\beta, c, a) F_5(\beta, -c, -a),$

$$U_{10} = \beta \{P^2 (P^2 - a^2 c^2 Q^2)^2 - Q^2 a^2 c^2 (P^2 - \beta^4 Q^2)^2\},$$
$$V_{10} = c \{P^2 (P^2 - \beta^4 Q^2)^2 - Q^2 \beta^4 (P^2 - a^2 c^2 Q^2)^2\}.$$

We have $F_5(\beta, a, c) F_{10}(\beta, a, c) = Q^3 F_5(\beta_1, a_1, c_1).$

$n = 11$ $F_{11}(\beta, a, c) = \beta (P^2 - a^2 c^2 Q^2)[(\beta^2 - c^2) P - c^2 (\beta^2 - a^2) Q]$

$$- c (P^2 - \beta^4 Q^2)[(\beta^2 - a^2) P - a^2 (\beta^2 - c^2) Q],$$
$$U_{11} = -F_{11}(\beta, c, a) F_{11}(\beta, -c, -a),$$
$$V_{11} = F_{11}(\beta, a, c) F_{11}(\beta, -a, -c).$$

Digitized by Google

$n = 12$ $F_{12}(\beta, a, c) = (P^3 - a^2c^2Q^2)(P + \beta^2 Q),$
$= - Q^3 F_6(\beta_1, a_1, c_1),$
$U_{12} = - P(M^2 - N^2Q^4\beta^4a^2c^2),$
$V_{12} = \beta c(P^2 - a^2c^2Q^2)(P^2 - \beta^4Q^2) H.$

$n = 13$ $F_{13}(\beta, a, c) = \beta(P^2 - a^2c^2Q^2)(P^3 - P^2Qc^2 - PQ^2\beta^2c^2 + Q^3\beta^2a^2c^2)$
$+ c(P^2 - \beta^4Q^2)(P^3 - P^2Qa^2 - PQ^2\beta^2a^2 + Q^3\beta^2a^2c^2),$
$U_{13} = F_{13}(\beta, c, a) F_{13}(\beta, - c, - a),$
$V_{13} = F_{13}(\beta, a, c) F_{13}(\beta, - a, - c).$

$n = 14$ $F_{14}(\beta, a, c) = F_7(\beta, - a, - c) F_7(\beta, c, a) F_7(\beta, - c, - a),$
$U_{14} = \beta\{(P^2 - a^2c^2Q^2)^2 M^2 - P^2Q^2a^2c^2(P^2 - \beta^4Q^2)^2 N^2\},$
$V_{14} = - c\{(P^2 - \beta^4Q^2)^2 M^2 - P^2Q^2\beta^4(P^2 - a^2c^2Q^2)^2 N^2\}.$
We have $F_7(\beta, a, c) F_{14}(\beta, a, c) = Q^6 F_7(\beta_1, a_1, c_1).$

$n = 15$ $F_{15}(\beta, a, c)$ is found from
$F_3(\beta, a, c) F_5(\beta, a, c) F_{15}(\beta, a, c)$
$= M(P^3 - P^2Qc^2 - PQ^2\beta^2c^2 + Q^3\beta^2a^2c^2)$
$- \beta cPQN(P^3 - P^2Qa^2 - PQ^2\beta^2a^2 + Q^3\beta^2a^2c^2).$
Denoting the right side by $\Phi_{15}(\beta, a, c)$, we have
$U_{15} = - \Phi_{15}(\beta, c, a) \Phi_{15}(\beta, - c, - a),$
$V_{15} = \Phi_{15}(\beta, a, c) \Phi_{15}(\beta, - a, - c).$

$n = 16$ $F_{16}(\beta, a, c) = - MN,$
$U_{16} = - (M^4 - P^4Q^4\beta^4a^2c^2N^4),$
$V_{16} = \beta cPQMN\{2M^2 - N^2(P^4 + Q^4\beta^4a^2c^2)\}.$
We have $F_{16}(\beta, a, c) = - Q^6 F_8(\beta_1, a_1, c_1).$

$n = 17$ $F_{17}(\beta, a, c) = M(\beta^2 - c^2)\{\beta^2(P^2 - a^2c^2Q^2)^2 - c^2(P - a^2Q)^2(P + \beta^2Q)^2\}$
$+ \beta cPQN(\beta^2 - a^2)\{\beta^2(P^2 - a^2c^2Q^2)^2 - a^2(P - c^2Q)^2(P + \beta^2Q)^2\},$
$U_{17} = F_{17}(\beta, c, a) F_{17}(\beta, - c, - a),$
$V_{17} = F_{17}(\beta, a, c) F_{17}(\beta, - a, - c).$

$n = 18$ $F_{18}(\beta, a, c) = F_9(\beta, - a, - c) F_9(\beta, c, a) F_9(\beta, - c, - a),$
$U_{18} = \beta(P^2 - a^2c^2Q^2)\{P^2(M^2 - N^2Q^4\beta^4a^2c^2)^2 - Q^2a^2c^2(P^2 - \beta^4Q^2)^4 H^2\},$
$V_{18} = c(P^2 - \beta^4Q^2)\{P^2(M^2 - N^2Q^4\beta^4a^2c^2)^2 - Q^2\beta^4(P^2 - a^2c^2Q^2)^4 H^2\}.$
We have $F_9(\beta, a, c) F_{18}(\beta, a, c) = Q^9 F_9(\beta_1, a_1, c_1).$

Digitized by Google

$n = 19$ $F_{19}(\beta, a, c) = U_9 V_{10} + V_9 U_{10}$,

$U_{19} = -F_{19}(\beta, c, a)\, F_{19}(\beta, -c, -a)$,

$V_{19} = F_{19}(\beta, a, c)\, F_{19}(\beta, -a, -c)$.

$n = 20$ $F_{20}(\beta, a, c) = -[P^2(P^2 - a^2c^2Q^2)^2 - Q^2a^2c^2(P^2 - \beta^4Q^2)^2]$
$\times [P(P^2 - \beta^4Q^2) - Q\beta^2(P^2 - a^2c^2Q^2)]$,

$U_{20} = -P\{M^2(M^2 - N^2Q^4\beta^4a^2c^2)^2$
$- N^2Q^4\beta^4a^2c^2(M^2 - P^4N^2)^2\}$,

$V_{20} = Q\beta c\{M^2(M^2 - P^4N^2)^2 - N^2P^4(M^2 - N^2Q^4\beta^4a^2c^2)^2\}$.

We have $F_{20}(\beta, a, c) = -Q^9 F_{10}(\beta_1, a_1, c_1)$.

$n = 24$ $F_{24}(\beta, a, c) = -(M^2 - N^2Q^4\beta^4a^2c^2)\, H$.

This is equal to $-Q^{12}F_{12}(\beta_1, a_1, c_1)$.

§ 3. As an illustration of the preceding formulae, we may consider polygons circumscribed about the circle

$$x^2 + y^2 = r^2,$$

with angular points lying on the circle

$$(x - d)^2 + y^2 = R^2.$$

The condition that the chord, of the latter, whose equation is

$$(x - d)\cos\tfrac{1}{2}(\alpha + \beta) + y\sin\tfrac{1}{2}(\alpha + \beta) = R\cos\tfrac{1}{2}(\alpha - \beta),$$

should touch the former is

$$a(1 + \phi^2\theta^2) + b\phi\theta + c(\phi^2 + \theta^2) = 0,$$

where, putting

$$p = \frac{R - d}{r}, \qquad q = \frac{R + d}{r}, \qquad k^2 = \frac{1 - p^2}{1 - q^2},$$

and

$$\theta = k^{\frac{1}{2}}\tan\tfrac{1}{2}\alpha, \qquad \phi = k^{\frac{1}{2}}\tan\tfrac{1}{2}\beta,$$

we have

$$a = \frac{1 - p^2}{k}, \qquad b = -2pq, \qquad c = 1.$$

Now put

$$s = \frac{1}{p} + \frac{1}{q} = \frac{2Rr}{R^2 - d^2},$$

$$t = \frac{4}{(s^2 - 1)pq} = \frac{4r^2(R^2 - d^2)}{4R^2r^2 - (R^2 - d^2)^2};$$

then we have

$$a^2 = \frac{s^2 - (4t^{-1} - 1)^2}{s^2 - 1}, \qquad c = 1, \qquad \beta = 2t^{-1} - 1.$$

The substitution of these values in the forms $F_n(\beta, a, c) = 0$ found above gives the conditions for closure in terms of s^2 and t, and

Digitized by Google

hence in terms of R, r, d; we have already said that $F_n(\beta, a, c)$ is an even function of a^2.

We may remark* that the interchange of R and d, in the resulting formula, changes the sign of p, and hence changes β, a, c into $-\beta, a, c$. Thus it changes $F_n(\beta, a, c) = 0$ into $F_n(\beta, -a, -c) = 0$. When n is odd, $F_n(\beta, -a, -c)$ is one of the three factors of $F_{2n}(\beta, a, c)$. See § 4 below.

The results for the conditions of closure in the cases $n = 3$ to $n = 8$ may be reduced to the following forms:

$n = 3$. Triangle. $\beta + c = 0$, or $t^{-1} = 0$, giving $s^2 - 1 = 0$.

$n = 4$. Quadrilateral. $R - d = 0$ or, alternatively, $t = 2$.

$n = 5$. Pentagon. $s^2(t-1)^2 - 1 = 0$.

$n = 6$. Hexagon. $(t-1)[s^2(t-1) + t - 3] = 0$.

$n = 7$. Heptagon. $[s(t-1) - s^{-1}]^2 - (t-1)^2(t-2)^2 = 0$.

$n = 8$. Octagon.

$$[s^2(t-1) - 1]\left[\frac{s^2(t-1) - 1}{s^2(t-1) + 1} + \frac{(t-1)(t-2)}{t}\right] = 0.$$

Of these two factors the first leads to

$$R^2 - d^2 = \pm 2r(\pm Rd)^{\frac{1}{2}}\dagger.$$

If we put $u = p^{-1} = r/(R-d)$, $v = q^{-1} = r/(R+d)$, so that $s = u + v$, $(s^2 - 1)t = 4uv$, we have

$$\beta : a : c = \tfrac{1}{2}(u^2 + v^2 - 1) : [(u^2 - 1)(v^2 - 1)]^{\frac{1}{2}} : uv,$$

and u^2, v^2 are the roots of the equation

$$\theta^2 - \theta\frac{a^2 - c^2 - 4\beta^2}{a^2 - c^2} + \frac{4\beta^2 c^2}{(a^2 - c^2)^2} = 0.$$

§ 4. The following statement of formulae will shew how the results in § 2 have been computed, and furnish various ways in which they can be verified.

Beside the polynomials $F_n(\beta, a, c)$, whose vanishing gives the reduced condition of closure, we may consider polynomials $\Phi_n(\beta, a, c)$, whose vanishing gives the complete condition of closure. Theoretically the former can be obtained from the latter by the equation

$$\Phi_n(\beta, a, c) = \prod_{d>2}^{n} F_d(\beta, a, c),$$

where d denotes all divisors of n greater than 2.

The polynomials Φ_n (and F_n), and the polynomials U_n, V_n are

* Cf. Chaundy, *Proc. Lond. Math. Soc.* XXII. 1924, p. 110.

† E. C. Titchmarsh, *Messenger of Math.* LII. 1922, p. 42.

Digitized by Google

deduced concurrently by the facts following, where A_n denotes U_n/V_n:

(*a*) $U_1 = V_1 = 1; \quad U_2 = \beta, \quad V_2 = c.$

(*b*) When n is odd, $= 2m + 1$,

$$\Phi_n(\beta, a, c) = (-1)^k (U_m V_{m+1} + V_m U_{m+1}),$$

where k is $\frac{1}{4}(n \pm 3)$ according as n is of the form $4s + 1$, or of the form $4s - 1$; namely, $(-1)^k = -(-1)^{\frac{1}{2}(n^2-1)}$.

(*c*) When n is even, $= 2m$,

$$\Phi_n(\beta, a, c) = (-1)^k U_m V_m,$$

where k is $\frac{1}{4}(n - 4)$ or $\frac{1}{4}(n - 2)$ according as n is of the form $4s$, or of the form $4s + 2$; namely, $(-1)^k = -(-1)^{\frac{1}{2}(n^2+2n)}$.

(*d*) When n is odd,

$$A_n = (-1)^{\frac{1}{2}(n-1)} \Phi_n(\beta, c, a)\Phi_n(\beta, -c, -a)/\Phi_n(\beta, a, c)\Phi_n(\beta, -a, -c).$$

(*e*) When n is even, $= 2m$,

$$A_n = \frac{\beta}{c} A_m(\beta_1, a_1, c_1).$$

A particular consequence is that, when n is odd,

$$F_{2n}(\beta, a, c) = F_n(\beta, -a, -c)\, F_n(\beta, c, a)\, F_n(\beta, -c, -a);$$

and it appears that

$$F_n(\beta, a, c)\, F_{2n}(\beta, a, c) = \pm Q^s F_n(\beta_1, a_1, c_1),$$

also for n odd ($\gtrless 3$), where s is determinable by the dimensions of the two sides. More generally, for any m, it appears that, with s similarly determined,

$$F_{4m}(\beta, a, c) = \pm Q^s F_{2m}(\beta_1, a_1, c_1).$$

We have, for $m > 1$,

$$V_{2m} = (-1)^m (U_{m-1} V_{m+1} - V_{m-1} U_{m+1})\, U_m V_m,$$
$$U_{2m} = (c^2 U^2_{m-1} U^2_{m+1} - a^2 V^2_{m-1} V^2_{m+1})\, \mu^{-1},$$

where μ is βc^2 or $a^2 - c^2$, according as m is odd or even.

Also $$U_m V_m = (U_{m-1} V_{m+1} + V_{m-1} U_{m+1})\, \sigma^{-1},$$

where σ is $-c$ or $(a^2 - c^2)/\beta c$, according as m is odd or even.

Also we have, the sign of $\Phi_{2m+1}(\beta, a, c)$ being such that it contains a term $\beta^{\frac{1}{2}m(m+1)}$ with positive sign,

$$U_{2m} - V_{2m} = (-1)^{m-1} \Phi_{2m-1}(\beta, a, c)\, \Phi_{2m+1}(\beta, -a, -c),$$
$$U_{2m} + V_{2m} = (-1)^{m-1} \Phi_{2m-1}(\beta, -a, -c)\, \Phi_{2m+1}(\beta, a, c),$$
$$U_{2m} + \frac{a}{c} V_{2m} = (-1)^{m-1} \Phi_{2m-1}(\beta, -c, -a)\, \Phi_{2m+1}(\beta, c, a),$$
$$U_{2m} - \frac{a}{c} V_{2m} = (-1)^{m-1} \Phi_{2m-1}(\beta, c, a)\, \Phi_{2m+1}(\beta, -c, -a).$$

Digitized by Google

These shew, incidentally, that U_{2m} and V_{2m}/c are both symmetrical in regard to a and c. But they have the great theoretical importance of shewing that U_{2m}, V_{2m} are both expressible by means of the irreducible polynomials F_n, in which n is odd. When n is odd, U_n, V_n have already been expressed by these irreducible polynomials. We may find formulae for $U_n \pm V_n$, $U_n \pm \frac{a}{c} V_n$ when n is odd; we have however expressed U_n and V_n each as a product of two polynomials F_n. It will be found (for n odd) that

$$U_n^2 - V_n^2 = - V_{n-1} V_{n+1} [(a^2 - c^2)/\beta c^2],$$

$$U_n^2 - \frac{a^2}{c^2} V_n^2 = U_{n-1} U_{n+1} [(a^2 - c^2)/\beta c^2].$$

It also appears, from the formula given above for $\Phi_{2n}(\beta, a, c)$, if $4m = 2^k N$, where N is odd and $k \geqslant 2$, that

$$\prod_{\delta=1}^{N} F_{2^k \delta}(\beta, a, c) = (-1)^{E\left(\frac{m+1}{2}\right)} U_{2m} (U_{m-1} V_{m+1} - V_{m-1} U_{m+1}),$$

where δ denotes all the divisors of N, and $E\left(\frac{m+1}{2}\right)$ is the greatest integer in $\frac{1}{2}(m+1)$. This formula determines $F_{4m}(\beta, a, c)$.

§ 5. The preceding formulae deal with algebraic questions. But they have in part been found in connexion with the theory of elliptic functions. An indication of this may be given.

There exists an odd doubly periodic function, $\phi(u)$, of the argument u, becoming infinite to the first order for $u = 0$, so that the limit of $u\phi(u) = 1$, which satisfies the differential equation

$$[\phi'(u)]^2 = [\phi(u)]^4 + \lambda [\phi(u)]^2 + h^4,$$

wherein λ, h^4 are constants, and $\phi'(u) = \partial\phi(u)/\partial u$. This equation shews that

$$\left(\frac{\partial^2}{\partial u^2} - \frac{\partial^2}{\partial v^2}\right) \log [\phi(u) + \phi(v)] = \phi^2(u) - \phi^2(v)$$
$$= \left(\frac{\partial^2}{\partial u^2} - \frac{\partial^2}{\partial v^2}\right) \log [\phi(u) - \phi(v)],$$

and, hence, that the function

$$\frac{1}{2}\left(\frac{\partial}{\partial u} + \frac{\partial}{\partial v}\right) \log \frac{\phi(u) + \phi(v)}{\phi(u) - \phi(v)}$$

is a function of $u + v$. This function is

$$\frac{\phi(u)\phi'(v) - \phi(v)\phi'(u)}{\phi^2(u) - \phi^2(v)}.$$

Digitized by Google

Hence, putting $v = 0$, we infer that this function is $\phi(u+v)$; and a second form for this is

$$\frac{h^4 - \phi^2(u)\phi^2(v)}{\phi(u)\phi'(v) + \phi(v)\phi'(u)}.$$

We can take the periods, Ω, Ω', of this function so that the second pole of $\phi(u)$ is for $u = \frac{1}{2}\Omega'$, the limit of $(u - \frac{1}{2}\Omega')\phi(u)$ being -1, and so that the zeros are for $u = \frac{1}{2}\Omega$, $u = \frac{1}{2}\Omega + \frac{1}{2}\Omega'$; further so that

$$\phi'(\tfrac{1}{2}\Omega), = -\phi'(\tfrac{1}{2}\Omega + \tfrac{1}{2}\Omega'), = h^2,$$

$$\phi(u + \tfrac{1}{2}\Omega) = h^2\phi(u), \qquad \phi(u + \tfrac{1}{2}\Omega') = -\phi(u).$$

The solutions of the equation $\phi(u) = \phi(v)$ are then

$$u \equiv v, \qquad u \equiv -v + \tfrac{1}{2}\Omega', \qquad (\text{mod. } \Omega, \Omega').$$

By writing the addition equation in the form

$$\phi_3(\phi_1^2 - \phi_2^2) - \phi_1\phi_2' = -\phi_2\phi_1',$$

and squaring, we can infer that, if we take

$$a = h^2, \quad b = -2\phi'(v), \quad c = -\phi^2(v), \quad \beta = -\phi(v)\phi(2v),$$

the (2, 2) relation

$$a(1 + \Phi^2\Theta^2) + b\Phi\Theta + c(\Phi^2 + \Theta^2) = 0$$

connects the two functions

$$\Theta, = \frac{1}{h}\phi(u), \qquad \Phi, = \frac{1}{h}\phi(u+v).$$

Further, that there is a relation

$$a_{n-1}(1 + \Theta_n^2\Theta^2) + b_{n-1}\Theta_n\Theta + c_{n-1}(\Theta_n^2 + \Theta^2) = 0,$$

in which

$$a_{n-1} = h^2, \qquad b_{n-1} = -2\phi'(nv), \qquad c_{n-1} = -\phi^2(nv),$$

connecting the two functions, Θ and

$$\Theta_n, = \frac{1}{h}\phi(u + nv).$$

The condition for closure after n links is then that

$$nv = P\Omega + P'\Omega',$$

where P, P' are integers. The reduced condition of closure $F_n(\beta, a, c) = 0$ is the relation connecting

$$\beta, = -\phi(v)\phi(2v), \qquad a = h^2, \qquad c = -\phi^2(v)$$

which holds for the values of v of the form $(P\Omega + P'\Omega')/n$, in which

Digitized by Google

P and P' have no common factor which divides n (and are not both zero).

The function A_n, or U_n/V_n, is $\phi(nv)/\phi(v)$. That this is a quotient of two rational homogeneous polynomials in β, a, c, follows by induction; for instance, from the formulae,

$$A_{n+1}A_{n-1} = -(A_n^2 - a^2/c^2)/(A_n^2 - 1),$$

$$A_{n+1} + A_{n-1} = (a^2 - c^2)\, A_n/\beta c\,(A_n^2 - 1),$$

for we have $A_1 = 1$, $A_2 = \beta/c$.

The same is true of $\phi'(v)/\phi^2(v)$, for we have

$$\phi'(v)/\phi^2(v) = \tfrac{1}{2}(A_{n-1} - A_{n+1})(A_n^2 - 1).$$

Also $\phi(nv)/\phi(v)$ is a rational function of β/c and a^2/c^2; if we allow the entry of the constant λ, these two are connected by the equation

$$4\left(\frac{\beta}{c}\right)^2\left[\frac{a^2}{c^2} - \frac{\lambda}{h^2}\frac{a}{c} + 1\right] = \left(\frac{a^2}{c^2} - 1\right)^2.$$

If we regard β, a, c as coordinates in a plane, the well-known theorems as to the algebraic solution of the multiplication equation become theorems for the algebraic determination of the intersections of the quartic curve represented by this last equation with the curves represented by the various equations $F_n(\beta, a, c) = 0$. For $n = 3, 5$ this last curve is of genus 0, 1 respectively; for $n = 7$ it can be shewn to be of genus 4.

Also, it may be remarked, when the odd number n is not a prime, the functions $F_n(\beta, a, c)$ can be expressed by functions $F_d(\beta_e, a_e, c_e)$, where $n = d(e+1)$, the $a_e, c_e, \beta_e, = (a_e^2 - c_e^2)/b_e$, being as above.

§ 6. It may be desirable to compare the formulae of this paper with those given by Halphen*. Suppose we require to inscribe polygons in a conic $S = 0$, whose sides shall touch a conic $\sigma = 0$. Let the discriminant of $\lambda S + \sigma = 0$ be written $\lambda^3\Delta + \lambda^2\Theta + \lambda\vartheta + \delta$. Halphen expresses the condition in terms of x, y given by

$$x = -(\vartheta^2 - 4\delta\Theta)^3/2^8\Delta^2\delta^4, \qquad y = -(\vartheta^3 - 4\vartheta\Theta\delta + 8\Delta\delta^2)/2^3\Delta\delta^2;$$

in terms of the s, t used above these may be taken to be

$$x = (s^2 - 1)\, s^{-4}t^{-2}, \qquad y = -(t - 2)\, s^{-2}t^{-1};$$

or, in terms of the β, a, c used above, if we put $\xi = x/y$, we may suppose

$$\xi = (\beta + c)^3/M, \qquad y = -\beta(a^2 - c^2)/M,$$

where

$$Mc^{-1} = (2\beta + c)^2 - a^2.$$

* *Fonct. Ellipt.* II. 1888, p. 377.

Digitized by Google

The explicit conditions of closure, for a polygon of 3, 4, ..., 11 sides, given by Halphen, are then, respectively,

(3) $\xi = 0$; (4) $y = 0$; (5) $\xi - 1 = 0$; (6) $\xi + y - 1 = 0$;

(7) $\xi(\xi - 1) + y = 0$; (8) $(\xi - 1)(2\xi - 1) + \xi y = 0$;

(9) $(\xi - 1)^3 - y(\xi + y - 1) = 0$;

(10) $\xi(\xi + y - 1)^2 + y(\xi^2 - \xi + y) = 0$;

(11) $\xi(\xi + y - 1)^3 + (\xi - 1)^3(\xi^2 - \xi + y) = 0$.

These forms, which Halphen shews (*loc. cit.*, p. 600) to agree with the more elegant forms given by Cayley*, do not bring out, for instance, that the condition for closure after an even number of links is a product of three factors, each rational in β, a, c, that is, in the coefficients of the (2, 2) relation. As an example the reduction of Halphen's condition for a hexagon may be carried out.

* *Papers*, II. pp. 87, 138.

Digitized by Google

Digitized by Google

Digitized by Google

PROCEEDINGS

OF THE

Cambridge Philosophical Society.

On a group of 1440 *birational transformations of four variables that arises in considering the projective equivalence of double sixes.*

By Mr W. BURNSIDE, Honorary Fellow of Pembroke College.

[*Received* 10 December 1925, *read* 18 January 1926.]

The lines of a double-six will here be represented by the usual notation

$$a,\ b,\ c,\ d,\ e,\ f,$$
$$a',\ b',\ c',\ d',\ e',\ f',$$

where two lines whose symbols are in the same line or same column of this scheme are non-intersectors and all other pairs of lines intersect. Any six of the lines, no two of whose symbols are in the same column, and just three are in the same row, are generators of a quadric, and the actual position in space of each of the other six is determined by the two points in which it intersects this quadric.

A point on the quadric is determined by the parameters of the two generating lines that pass through it, and these parameters may be chosen so that to three arbitrarily assigned generators of either system the values 0, 1, ∞ of the parameter correspond.

Hence if S is the quadric of which a, b, c, d', e', f' are generators, the parameters λ, μ of the generators of S may be chosen so that:

$$a,\ b,\ c \text{ are } \lambda = 0,\ \lambda = 1,\ \lambda = \infty,$$

$$d',\ e',\ f' \text{ are } \mu = 0,\ \mu = 1,\ \mu = \infty,$$

a' meets S where $\lambda = 1,\ \mu = u_1$ and $\lambda = \infty,\ \mu = U_1$,

b' ,, ,, $\lambda = \infty,\ \mu = u_2$,, $\lambda = 0,\ \mu = U_2$,

c' ,, ,, $\lambda = 0,\ \mu = u_3$,, $\lambda = 1,\ \mu = U_3$,

d ,, ,, $\lambda = v_1,\ \mu = 1$,, $\lambda = V_1,\ \mu = \infty$,

e ,, ,, $\lambda = v_2,\ \mu = \infty$,, $\lambda = V_2,\ \mu = 0$,

f ,, ,, $\lambda = v_3,\ \mu = 0$,, $\lambda = V_3,\ \mu = 1$.

Now each of the lines a', b', c' meets each of the lines d, e, f. This gives nine equations connecting the u's, U's, v's and V's. It

Digitized by Google

is readily found that one of these equations is a consequence of the other eight, and that these give

$$u_3 = U_1 U_2 \cdot \frac{(U_1-1)(U_2-1)-(u_1-1)(u_2-1)}{u_1u_2(U_1-1)(U_2-1)-U_1U_2(u_1-1)(u_2-1)},$$

$$U_3 = u_1u_2 \cdot \frac{(U_1-1)(U_2-1)-(u_1-1)(u_2-1)}{u_1u_2(U_1-1)(U_2-1)-U_1U_2(u_1-1)(u_2-1)},$$

$$v_1 = \frac{(U_1-u_1)(U_2-1)}{(U_1-1)(U_2-1)-(u_1-1)(u_2-1)},$$

$$V_1 = \frac{(U_1-u_1)(u_2-1)}{(U_1-1)(U_2-1)-(u_1-1)(u_2-1)},$$

$$v_2 = \frac{(U_1-u_1)u_2}{U_1U_2-u_1u_2},$$

$$V_2 = \frac{(U_1-u_1)U_2}{U_1U_2-u_1u_2},$$

$$v_3 = \frac{(U_1-u_1)U_2(u_2-1)}{u_1u_2(U_1-1)(U_2-1)-U_1U_2(u_1-1)(u_2-1)},$$

$$V_3 = \frac{(U_1-u_1)(U_2-1)u_2}{u_1u_2(U_1-1)(U_2-1)-U_1U_2(u_1-1)(u_2-1)}.$$

This is, of course, only one form of solution of the nine equations. It corresponds to the supposition that besides a, b, c, d', e', f' the two lines a', b' are taken arbitrarily. There will be five other forms of solution corresponding to the suppositions that b', c'; c', a'; d, e; e, f; or f, d respectively are taken arbitrarily. Since the double-six is determined uniquely when either b', c' or e, f are given in addition to a, b, c, d', e', f', the last four of the above equations must determine u_1, U_1, u_2, U_2 uniquely in terms of v_2, V_2, v_3, V_3.

If the notation ab' (sequence of symbols immaterial) is used to denote the point of intersection of the lines a and b', the values of μ at the four points $d'b, e'b, f'b, a'b$ are $0, 1, \infty, u_1$; so that u_1 is the cross-ratio of the two pairs of points $a'b, e'b$; $d'b, f'b$. The symbols U_1, u_2 and U_2 have similar significations. Hence when u_1, U_1, u_2, U_2 are given, the cross-ratio of each such two pairs of points on any line of the double-six is uniquely determined.

Now consider the permutation

$$\begin{pmatrix} a, & b, & c, & d, & e, & f, & a', & b', & c', & d', & e', & f' \\ a_1, & b_1, & c_1, & d_1, & e_1, & f_1, & a_1', & b_1', & c_1', & d_1', & e_1', & f_1' \end{pmatrix}$$

of the twelve symbols, where $a_1, b_1, c_1, d_1, e_1, f_1$ are a, b, c, d, e, f in some altered sequence: $a_1', b_1', c_1', d_1', e_1', f_1'$ are a', b', c', d', e', f' in an altered sequence; and

$$\begin{pmatrix} a', & b', & c', & d', & e', & f' \\ a_1', & b_1', & c_1', & d_1', & e_1', & f_1' \end{pmatrix}$$

Digitized by Google

gives the same permutation of a', b', c', d', e', f' that

$$\begin{pmatrix} a, & b, & c, & d, & e, & f \\ a_1, & b_1, & c_1, & d_1, & e_1, & f_1 \end{pmatrix}$$

gives of a, b, c, d, e, f.

Let S_1 be the quadric containing $a_1, b_1, c_1, d_1', e_1', f_1'$. The parameters of its generators may be chosen so that

$$a_1, \ b_1, \ c_1 \text{ are } \lambda = 0, \ \lambda = 1, \ \lambda = \infty,$$
$$d_1', \ e_1', \ f_1' \text{ are } \mu = 0, \ \mu = 1, \ \mu = \infty.$$

Then if

a_1' meets S_1 where $\lambda = 1, \quad \mu = u_1'$ and $\lambda = \infty, \ \mu = U_1'$,
and b_1' ,, S_1 ,, $\lambda = \infty, \ \mu = u_2' \quad$,, $\lambda = 0, \quad \mu = U_2'$,
u_1', U_1', u_2', U_2' are rational functions of u_1, U_1, u_2, U_2. Also since the double-six is uniquely determined by u_1', U_1', u_2', U_2', it follows that u_1, U_1, u_2, U_2 can be uniquely determined in terms of u_1', U_1', u_2', U_2'.

Hence the equations

$$u_1' = f(u_1, U_1, u_2, U_2), \qquad u_2' = g(u_1, U_1, u_2, U_2),$$
$$U_1' = F(u_1, U_1, u_2, U_2), \qquad U_2' = G(u_1, U_1, u_2, U_2),$$

specify a birational substitution of the four symbols; and each permutation of the twelve original symbols of the kind specified gives rise to such a birational substitution on u_1, U_1, u_2, U_2. Moreover it is clearly the case that to the product of two of the permutations there corresponds the product of the two birational substitutions to which they give rise.

Hence the set of 720 birational substitutions is a representation (which has not hitherto been recognised) of the symmetric group of degree six.

There is some interest in shewing how the generating substitutions of the group in this form may be determined. The specifica-

	a $(\lambda=0)$	b $(\lambda=1)$	c $(\lambda=\infty)$	d	e	f
a'	×	$1, u_1$	∞, U_1	—	—	—
b'	$0, U_2$	×	∞, u_2	—	—	—
c'	$0, u_3$	$1, U_3$	×	—	—	—
d' $(\mu=0)$	0, 0	1, 0	$\infty, 0$	×	$V_2, 0$	$v_2, 0$
e' $(\mu=1)$	0, 1	1, 1	$\infty, 1$	$v_1, 1$	×	$V_3, 1$
f' $(\mu=\infty)$	$0, \infty$	$1, \infty$	∞, ∞	V_1, ∞	v_2, ∞	×

tion of the twelve u's, U's, v's, V's originally introduced is conveniently given by the diagram, where the symbols written in each compartment give the values of λ, μ at the corresponding point.

Digitized by Google

If the quadric S', determined by a', b', c', d, e, f, is used in the place of the quadric S, the diagram would be

	a	b	c	d $(\mu=0)$	e $(\mu=1)$	f $(\mu=\infty)$
a' $(\lambda=0)$	×	$0, U_2''$	$0, u_3''$	0, 0	0, 1	$0, \infty$
b' $(\lambda=1)$	$1, u_1''$	×	$1, U_3''$	1, 0	1, 1	$1, \infty$
c' $(\lambda=\infty)$	∞, U_1''	∞, u_2''	×	$\infty, 0$	$\infty, 1$	∞, ∞
d'	—	—	—	×	$v_1'', 1$	V_1'', ∞
e'	—	—	—	$V_2'', 0$	×	v_2'', ∞
f'	—	—	—	$v_3'', 0$	$V_3'', 1$	×

Now it is a known property of a double-six that there is a quadric, with respect to which a, a'; b, b'; ...; f, f' are polar lines. Reciprocation with respect to this replaces the four points

$$b, a', \quad d, a', \quad e, a', \quad f, a'$$

by the four planes

$$[b', a], \ [d', a], \ [e', a], \ [f', a],$$

where $[b', a]$ denotes the plane containing b' and a. These four planes are met by the line c in the four points

$$b', c, \quad d', c, \quad e', c, \quad f', c,$$

so that this set of points is projective with the set

$$b, a', \quad d, a', \quad e, a', \quad f, a'.$$

With the notation here used this gives

$$U_2'' = u_2,$$

and in the same way the relations

$$u_2'' = U_2, \quad U_1'' = u_1, \quad u_1'' = U_1, \quad \ldots \quad V_1'' = v_1, \quad v_1'' = V_1, \quad \ldots$$

are proved. Hence the second of the above diagrams is

	a	b	c	d	e	f
a'	×	$0, u_2$	$0, U_3$	0, 0	0, 1	$0, \infty$
b'	$1, U_1$	×	$1, u_3$	1, 0	1, 1	$1, \infty$
c'	∞, u_1	∞, U_2	×	$\infty, 0$	$\infty, 1$	∞, ∞
d'	—	—	—	×	$V_1, 1$	v_1, ∞
e'	—	—	—	$v_2, 0$	×	V_2, ∞
f'	—	—	—	$V_3, 0$	$v_3, 1$	×

Digitized by Google

Part of the diagram, corresponding to the permutation

$$\begin{pmatrix} a, & b, & c, & d, & e, & f, & a', & b', & c', & d', & e', & f' \\ b, & c, & d, & e, & f, & a, & b', & c', & d', & e', & f', & a' \end{pmatrix},$$

is

	a	b	c	d	e	f
a'	×	$0, \infty$	$1, \infty$	∞, ∞	—	—
b'	—	×	$1, u_1'$	∞, U_1'	—	—
c'	—	$0, U_2'$	×	∞, u_2'	—	—
d'	—	—	—	×	—	—
e'	—	$0, 0$	$1, 0$	$\infty, 0$	×	—
f'	—	$0, 1$	$1, 1$	$\infty, 1$	—	×

Comparing this with the two that have been obtained above, it follows that

$$\left.\begin{matrix} u_1', 0, 1, \infty \\ u_2, 1, \infty_1, U_1 \end{matrix}\right\}, \quad \left.\begin{matrix} U_1', 0, 1, \infty \\ 1, v_2, V_3, 0 \end{matrix}\right\}, \quad \left.\begin{matrix} u_2', 0, 1, \infty \\ \infty, v_2, V_3, 0 \end{matrix}\right\}, \quad \left.\begin{matrix} U_2', 0, 1, \infty \\ U_3, 1, \infty, u_1 \end{matrix}\right\}$$

are respectively projective ranges.

Hence

$$u_1' = \frac{u_2 - 1}{u_2 - U_1}, \quad U_1' = \frac{V_3(1 - v_2)}{(V_3 - v_2)}, \quad u_2' = \frac{V_3}{V_3 - v_2}, \quad U_2' = \frac{U_3 - 1}{U_3 - u_1}.$$

When the values of v_2, V_3 and U_3 are entered in this it becomes

$$\left.\begin{aligned} u_1' &= \frac{u_2 - 1}{u_2 - U_1} \\ U_1' &= \frac{U_2 - 1}{U_2 - u_1} \\ u_2' &= \frac{(U_2 - 1)(U_1 U_2 - u_1 u_2)}{U_1 (U_2 - u_1)(U_2 - u_2)} \\ U_2' &= \frac{(u_2 - 1)(U_1 U_2 - u_1 u_2)}{u_1 (u_2 - V_1)(U_2 - u_2)} \end{aligned}\right\} \quad \text{.....(A).}$$

The substitution that corresponds to the permutation

$$\begin{pmatrix} a, & b, & c, & d, & e, & f, & a', & b', & c', & d', & e', & f' \\ b, & a, & c, & d, & e, & f, & b', & a', & c', & d', & e', & f' \end{pmatrix}$$

is found readily in the form

$$u_1' = U_2, \quad U_1' = u_2, \quad u_2' = U_1, \quad U_2' = u_1 \quad \text{.....(B);}$$

and this, together with the substitution just obtained, generate

Digitized by Google

the group of 720 birational substitutions on four symbols which gives a representation of the symmetric group of degree six.

It has been seen above that, to the permutation,

$$\begin{pmatrix} a, & b, & c, & d, & e, & f, & a', & b', & c', & d', & e', & f' \\ a', & b', & c', & d', & e', & f', & a, & b, & c, & d, & e, & f \end{pmatrix}$$

of the twelve symbols, which does not occur among the 720 permutations already considered, and which is permutable with every one of them, corresponds the substitution

$$u_1' = U_1, \quad U_1' = u_1, \quad u_2' = U_2, \quad U_2' = u_2 \quad \ldots\ldots(C),$$

on the four symbols. Hence every substitution of the group of birational substitutions generated by (A) and (B) is permutable with the substitution (C); and (A), (B) and (C) generate a group of order 1440, which is simply isomorphic with the direct product of the symmetric group of degree six and an operation of order two; and in it (C) is a self-conjugate substitution of order two.

Suppose now that A, B, C are any three non-intersecting lines and D', E', F' any three lines which meet each of A, B, C.

Two arbitrarily given double-sixes can, by suitable collineations, be changed into two, six of whose lines are A, B, C, D', E', F'. These two are completely defined by the values

$$u_1, \quad U_1, \quad u_2, \quad U_2$$

and

$$u_1', \quad U_1', \quad u_2', \quad U_2'$$

of the parameters that have been introduced. Then the necessary and sufficient condition that the two double-sixes should be projectively equivalent is that u_1', U_1', u_2', U_2' and u_1, U_1, u_2, U_2 should be connected by one of the substitutions of the group generated by (A), (B) and (C).

Digitized by Google

The General (m, n) Correspondence. By R. VAIDYANATHASWAMY. (Communicated by Mr H. W. TURNBULL.)

[*Received* 28 January, *read* 1 March 1926.]

CONTENTS.

§ 1. THE GENERAL (*m*, *n*) CORRESPONDENCE.

(1) Double Binary Forms in independent and in cogredient variables.
(2) Canonical Forms of correspondences with given fixed points.
(3) Multipliers.
(4) Defective Forms.
(5) Derivate Forms.

§ 2. COMPLETE AND CLOSED SETS OF A CORRESPONDENCE.

(6) Canonical Forms with Complete Sets.
(7) Reduction of Rank through Complete Set.
(8) Geometrical Interpretation in *n* Dimensions.
(9) Complete and Closed *n*-set for $(n+1, n)$ Forms.

INTRODUCTION.

The present work is the outcome of a study of the special type of (3, 2) correspondence on a circle, namely that between a point and the extremities of its pedal line. This led to considering correspondences in general, and to the formulation of concepts which are believed to be new, for example the Canonical Forms, § 1 (2), and Multipliers, § 1 (3). The case of the $(n, 1)$ correspondence has already been given* by the author.

In § 1 (4) *Wertigkeit* (cf. Waelsch, "Uber Binaren Formen...," *Monatshefte für Math. und Physik*, Wien (6) 1895) has been rendered by *defect*. The phrase *defective polar forms* has been adopted for convenience.

§ 1. THE GENERAL (*m*, *n*) CORRESPONDENCE.

(1) *Double-binary Forms in independent and in cogredient variables.*

There is bound to be a great difference in the algebraic treatment of double-binary forms in digredient variables x, y and those in cogredient variables; for, though theoretically the only difference is in the introduction of a new invariant type (xy), yet the possibility of "fixed elements" peculiar to forms in cogredient variables—or, in other words, to correspondences between the elements of the same binary field—compels certain lines of procedure, which have nothing similar to them in the case of forms in

* Cf. "Theory of the Rational Transformation", *Journal of the Indian Math. Soc.* (1921).

Digitized by Google

independent variables, or correspondences between different binary fields. While this is so, it is true on the other hand that all properties of correspondences between distinct binary fields must continue to hold when the fields become identical; for since the change from digredience to cogredience only replaces the group of transformations of the variables by one of its own subgroups, it follows that any general theory of forms in independent variables cannot become invalidated by the variables becoming cogredient. It would thus appear that the forms in cogredient variables, while being less general, must also be considerably richer than those in independent variables.

As an important instance of a body of general ideas continuing to be valid when the variables become cogredient, we may mention the rank-theory* of the double-binary form. An (m, n) form $F(x, y)$ is said to be of rank r, when it can be expressed in the shape

$$F(x, y) = f_1(x)\,\phi_1(y) + f_2(x)\,\phi_2(y) + \ldots + f_r(x)\,\phi_r(y),$$

but not in the shape $f_1(x)\,\phi_1(y) + \ldots + f_{r-1}(x)\,\phi_{r-1}(y)$. In particular, F is of rank one if it is a product of binary forms. The necessary and sufficient condition that F may be of rank r is that G_r be the first of the rank-covariants $G_1, G_2, \ldots$ which vanishes identically; where

$$F(x, y) = a_x^m b_y^n = a_{px}^m b_{py}^n \quad (p = 0, 1, 2, \ldots),$$

$$G_r = (a_0 a_1) \ldots (a_0 a_r)(a_1 a_r) \ldots (a_{r-1} a_r)(b_0 b_1) \ldots \times (b_{r-1} b_r)\, a_{0x}^{m-r} \ldots a_{rx}^{m-r} b_{0y}^{n-r} \ldots b_{ry}^{n-r}.$$

In particular, G_1 is the $(1, 1)$ transvectant of F. If F is of rank r, and is expressed in the shape

$$F(x, y) = f_1(x)\,\phi_1(y) + \ldots + f_r(x)\,\phi_r(y),$$

it may be shewn that the $(k-1)$th rank-covariant G_{k-1} is given by

$$G_{k-1} = \Sigma\, J_{12\ldots k}(x)\, J'_{12\ldots k}(y),$$

where $J_{12\ldots k}$ and $J'_{12\ldots k}$ are the Jacobians† of the forms $f_1 f_2 \ldots f_k$ and $\phi_1 \phi_2 \ldots \phi_k$ respectively. In particular, the last non-vanishing rank-covariant G_{r-1} is the product of the Jacobians of all the forms

* Cf. *Proc. London Math. Soc.* 2, 24 (1925), p. 83.

† The Jacobian of $r+1$ binary m-ics $f_k(x) = a_{kx}^m$ $(k = 0, 1, \ldots r)$ is symbolically

$$(a_0 a_1)(a_0 a_2) \ldots (a_0 a_r) \ldots (a_{r-1} a_r)\, a_{0x}^{m-r} \ldots a_{kx}^{m-r}$$

which is a numerical multiple of the determinant

$$\left| \frac{\partial^r f_0}{\partial x_1^r} \frac{\partial^r f_1}{\partial x_1^{r-1} \partial x_2} \cdots \frac{\partial^r f_r}{\partial x_2^r} \right|;$$

a combinant of the linear system

$$\lambda_0 f_0 + \lambda_1 f_1 + \ldots + \lambda_r f_r.$$

Digitized by Google

f and ϕ respectively, that is, of the Jacobians of the two linear systems

$$L_1: \lambda_1 f_1(x) + \lambda_2 f_2(x) + \ldots + \lambda_r f_r(x),$$
$$L_2: \mu_1 \phi_1(y) + \mu_2 \phi_2(y) + \ldots + \mu_r \phi_r(y).$$

Here L_1 is the system of forms of the type $(bc)^n a_x^m$, and L_2 of those of the type $(ad)^m b_y^n$, c_y^n and d_x^m being arbitrary binary forms. The systems L_1 and L_2 may be called the x- and the *y-concurrency systems* of F. If $m = n$, and F has its maximum rank $n + 1$, then both the concurrency systems are non-existent, that is to say, they comprise all possible forms. If $m > n$, and F has its maximum rank $n + 1$, then L_2 is non-existent and the Jacobian of L_1 is the binary form G_n of order $(m - n)(n + 1)$. In the particular case $m = n + 1$, G_n will be of order $n + 1 = m$, and therefore L_1 will be the system of $(n + 1)$-ics apolar to G_n. Hence:

1·1 *The concurrency system of an $(n + 1, n)$ form F of maximum rank is the system of $(n + 1)$-ics apolar to its last rank covariant G_n.*

All the above holds, mutatis mutandis, when the variables x, y become cogredient.

We shall be concerned throughout this paper only with forms in cogredient variables. We shall whenever possible make no distinction between forms and the correspondences obtained by equating them to zero; for instance, we shall denote the form and the correspondence by the same symbol, and in general speak of them interchangeably, whenever that will cause no error.

(2) *Canonical Forms of correspondences with given fixed points.*

An (m, n) correspondence $F(x, y) = 0$ has $m + n$ fixed points given by $F(x, x) = 0$. The form $F(x, y)$ involves $(m + 1)(n + 1)$ parameters linearly; if its fixed points are known, it will still involve $(m + 1)(n + 1) - (m + n) = mn + 1$ parameters linearly. To express $F(x, y)$ when its fixed points are known, we partition the $m + n$ fixed points (supposed distinct) in any way into two groups of m and n—$\alpha_1 \alpha_2 \ldots \alpha_m$ and $\beta_1 \beta_2 \ldots \beta_n$ say. The form $F(x, y)$ can then be expressed in only one way in the shape:

1·2

$$F(x, y) = \left\{\mu + \Sigma \mu_{pq} \frac{(\alpha_p y)(\beta_q x)}{(\alpha_p x)(\beta_q y)}\right\} (\alpha_1 x) \ldots (\alpha_m x)(\beta_1 y) \ldots (\beta_n y);$$
$$[p = 1, 2, \ldots m;\ q = 1, 2, \ldots n].$$

For this shape involves $mn + 1$ parameters μ, μ_{pq}. And when y is put equal to x, it becomes

$$(\mu + \Sigma \mu_{pq})(\alpha_1 x) \ldots (\alpha_m x)(\beta_1 x) \ldots (\beta_n x),$$

Digitized by Google

shewing that it has the given fixed points. Also, there are no values of μ, μ_{pq}, other than zero, for which the right side is identically zero; for, assuming such an identity and putting $x = \alpha_p$, $y = \beta_q$ in it, we have $\mu_{pq} = 0$ for all values of p, q, and therefore also $\mu = 0$. Thus the $mn + 1$ parameters μ, μ_{pq} are effective, and $F(x, y)$ admits of unique expression in this canonical shape. For different partitions of the fixed points, we get in general different canonical shapes; if however m or n is unity, we get the same canonical shape for all partitions.

On putting $x = \alpha_p$, $y = \beta_q$ in 1·2, we obtain the value of μ_{pq}, namely,

1·3
$$\mu_{pq} = \frac{-F(\alpha_p, \beta_q)}{(\alpha_p\beta_q)^2(\alpha_1\alpha_p)\dots(\alpha_{p-1}\alpha_p)(\alpha_{p+1}\alpha_p)\dots(\alpha_m\alpha_p)(\beta_1\beta_q)\dots(\beta_{q-1}\beta_q)(\beta_{q+1}\beta_q)\dots(\beta_n\beta_q)}$$

Further, if
$$F(x, x) = M(\alpha_1 x)\dots(\alpha_m x)(\beta_1 x)\dots(\beta_n x),$$
we have

1·4
$$\mu + \Sigma\mu_{pq} = M.$$

These equations determine all the parameters μ, μ_{pq} in terms of the fixed points and the coefficients of F. We shall find it convenient to write

1·5
$$\mu_{1q} + \mu_{2q} + \dots + \mu_{mq} = \mu_q',$$
$$\mu_{p1} + \mu_{p2} + \dots + \mu_{pn} = \mu_p,$$
so that
$$\sum_p \mu_p = \sum_q \mu_q' = M - \mu.$$

(3) *Multipliers.*

If we consider x to be the independent variable, and y the dependent variable, the correspondence F also represents a transformation from x to y. Supposing for the moment that x, y are non-homogeneous variables, we define the multiplier of F at a fixed point α to be the value of $\frac{dy}{dx}$ at the place $x = \alpha$, $y = \alpha$, of the associated Riemann surface. To obtain the value of the multiplier at the fixed point α_1 when the variables are homogeneous, we put $x = \alpha_1 + \delta z$, $y = \alpha_1 + \delta' z$ (z being an arbitrary point and δ, δ' infinitesimal scalars), and suppose that $F(x, y) = 0$ for these values. The limiting value of δ'/δ is then the required multiplier. Thus from 1·2, we have

$$\{\mu\delta + \Sigma_q \mu_{1q}\delta' + \Sigma_q \mu_{2q}\delta + \dots + \Sigma_q \mu_{mq}\delta\} \times (\alpha_1 z)(\alpha_2\alpha_1)\dots(\alpha_m\alpha_1)(\beta_1\alpha_1)\dots(\beta_n\alpha_1) = 0.$$

Digitized by Google

Hence ν_p, the multiplier at α_p, and ν_q', the multiplier at β_2, are given by

1·6 $$\nu_p = \frac{\mu_p - M}{\mu_p}; \quad \nu_q' = \frac{\mu_q'}{\mu_q' - M} \quad (1\cdot4,\ 1\cdot5).$$

We have thus the fundamental identity between the multipliers

$$\sum_{p=1}^{m} \frac{1}{1-\nu_p} + \sum_{q=1}^{n} \frac{\nu_q'}{1-\nu_q'} = 0.$$

This identity may be written in the symmetric shape

1·7 $$\Sigma \frac{1}{1-\nu} = n,$$

where the summation on the left is for all the multipliers, and the quantity n on the right is the order of F in the *dependent* variable.

If the form F is of rank one, say

$$F(x, y) = (\alpha_1 x) \dots (\alpha_m x)(\beta_1 y) \dots (\beta_n y),$$

then its canonical shape 1·2 for the partition $\alpha_1\alpha_2 \dots \alpha_m/\beta_1\beta_2 \dots \beta_n$ has only one term, and 1·6 shows that the multipliers at $\alpha_1\alpha_2 \dots \alpha_m$ become infinite, and that the remaining multipliers vanish. Conversely, if all the multipliers of F are infinite or zero, then by 1·7 exactly m must be infinite and n zero, and F is necessarily of rank one.

The *polar* (m, n) forms are those which result on polarising a binary $(m+n)$-ic in x n times with respect to y. On putting $y = x$ in a polar form, we simply revert to the binary form from which it was derived.

1·8 *The multipliers of the polar correspondence are all equal, each being equal (by* 1·7*) to* $-m/n$.

For, supposing that the fixed points $\alpha_1\alpha_2 \dots \alpha_{m+n}$ are all distinct, let us seek the condition that $m+n$ points $p_1 p_2 \dots p_{m+n}$ all infinitely near α_1 may form a group apolar to $(\alpha_1\alpha_2 \dots \alpha_{m+n})$. Relatively to the scale of the mutual distances of $\alpha_1 p_1 p_2 \dots p_{m+n}$ the points $\alpha_2\alpha_3 \dots \alpha_{m+n}$ are all infinitely far away. Hence so far as geometry in the immediate neighbourhood of α_1 is concerned, these latter points may all be considered to be coincident with the point at infinity on the line. Thus, the required condition of apolarity is simply that α_1 should be the centroid of $p_1 p_2 \dots p_{m+n}$. In particular, if x, y are in the neighbourhood of α_1, and are correspondents in the polar (m, n) correspondence defined by

$$(\alpha_1 x)(\alpha_2 x) \dots (\alpha_{m+n} x),$$

then the group composed of m points at x and n points at y is apolar to $(\alpha_1\alpha_2 \dots \alpha_{m+n})$; so that α_1 is the centroid of masses m, n at x, y respectively. Thus the multiplier of the polar correspondence at α_1 (and similarly at any other fixed point) is $-m/n$.

Digitized by Google

(4) *Defective Forms.*

If a correspondence F is such that $F(x, x) = 0$ identically (so that $F(x, y)$ has the factor (xy)), then the fixed points of F are indeterminate, and so F is an *identical*—or, as we shall say, a *defective*—correspondence. Among the (1, 1) correspondences there is only one type of identical correspondence, namely $(xy) = 0$; but for the (m, n) correspondences F there are different types according to the highest power of (xy) that divides F.

DEF. A form F is said to be of defect i, if it has the factor $(xy)^i$, but not the factor $(xy)^{i+1}$.

The condition that F, when expressed in the canonical shape 1·2, may be defective is $M = \mu + \Sigma\mu_{pq} = 0$; that is to say, a form with given fixed points has to satisfy only one condition for being defective. We see, as a matter of fact, that the $(mn + 1)$-parameter-family of (m, n) forms with given fixed points contains the mn-parameter-family consisting of all defective forms.

A polar form $P(x, y)$ is never defective; for, on putting $y = x$ in P, we revert to the binary form from which it was derived. This binary form not being supposed to vanish identically, it follows that P is not defective. We can therefore give a new meaning to the term "defective polar form." We accordingly define:

1·9 *The product of $(xy)^i$ and a polar form will be called a defective polar form, of defect i.*

Just as we have the invariant factor (xy) for cogredient variables, we also have the correlative invariant operator

$$\Omega = \frac{d}{dx_1}\frac{d}{dy_2} - \frac{d}{dx_2}\frac{d}{dy_1}.$$

The result of operating with Ω on any symbolic product is the sum of terms obtained by convolving every factor of the type (ax) with every factor of the type (βy) (that is, replacing $(ax)(\beta y)$ by $(a\beta)$).

1·10 The result of operating with Ω on a polar form is zero.

For a polar form being a symbolic product of the shape $(ax)^m (ay)^n$ every convolution in it leads to zero.

1·11 The result of operating with Ω on a defective polar (m, n) form $(xy)^i P(x, y)$ is

$$i(m + n - i + 1)(xy)^{i-1} P(x, y)^*.$$

Associated with the existence of the invariant factor (xy) and the correlative invariant operator Ω, there is a canonical shape, in

* Cf. Grace and Young, *Algebra of Invariants*, pp. 53–54 (1903).

Digitized by Google

which any form F can be expressed uniquely as the sum of a polar form and defective polar forms, namely:

$$1\cdot12 \quad F(x, y) = (P^{m+n})_{y^n} + (xy)(P^{m+n-2})_{y^{n-1}} + \ldots + (xy)^i (P^{m+n-2i})_{y^{n-i}} + \ldots,$$

where P^r is a binary form of order r in x and the suffix y^k denotes the kth polar with respect to y. The validity and uniqueness of this shape follow on observing that the number of parameters on both sides is the same. The form F is thus specified by the "ladder"*

$$(P^{m+n}, P^{m+n-2}, \ldots P^{m+n-2i}, \ldots).$$

The forms of the ladder can all be obtained by means of the operator Ω (1·11). Thus

$$1\cdot13 \quad [\Omega^i F(x, y)]_{y=x} \equiv i!\,(m+n-i+1)(m+n-i)\ldots(m+n-2i+2)\,P^{m+n-2i}.$$

(5) *Derivate Forms.*

If a pencil of (m, n) forms $\lambda F_1 + \mu F_2$ contains a defective member, it follows on putting $y = x$, that each member of the pencil has the same fixed points. Conversely:

1·14 *A pencil of (m, n) forms determined by two forms with the same fixed points necessarily contains a defective member.*

Now, given an (m, n) form F, there are $\binom{m+n}{n}$ forms of rank one with the same fixed points as F, each corresponding to a partition of the fixed points. [Thus, if $\alpha_1\alpha_2 \ldots \alpha_m/\beta_1\beta_2 \ldots \beta_n$ be a partition of the fixed points, the corresponding form of rank one is $(\alpha_1 x)\ldots(\alpha_m x)(\beta_1 y)\ldots(\beta_n y)$.] By 1·14 the pencil

$$F - \lambda(\alpha_1 x)\ldots(\alpha_m x)(\beta_1 y)\ldots(\beta_n y)$$

contains a defective member, $(xy)\,\phi(x, y)$ say.

1·15 We define the $(m-1, n-1)$ form $\phi_{\alpha}^{\beta}(x, y)$ to be the derivate of F in respect of the partition $\alpha_1\alpha_2 \ldots \alpha_m/\beta_1\beta_2 \ldots \beta_n$ of its fixed points.

The fundamental property of the derivate is clear from its definition, namely:

1·16 If ϕ be the derivate of F in respect of

$$\alpha_1\alpha_2 \ldots \alpha_m/\beta_1\beta_2 \ldots \beta_n,$$

then the correspondents of $x = \alpha_p$ in ϕ are all the correspondents of $x = \alpha_p$ in F, with the exception of α_p itself $(p = 1, 2, \ldots m)$; a similar statement holds about the correspondents of $y = \beta_q$ in ϕ $(q = 1, 2, \ldots n)$.

* Cf. Waelsch, *loc. cit.*

Digitized by Google

In particular, if $(\alpha_1\alpha_2 \dots \alpha_m)$ or $(\beta_1\beta_2 \dots \beta_n)$ be a complete set of F, it will be a semi-complete set of ϕ (see below, § 2 (6)).

Of special interest is the derivate of a form of rank one, in respect of a partition not its own. Consider for instance the derivate of $F = (\alpha_1 x) \dots (\alpha_m x)(\beta_1 y) \dots (\beta_n y)$ in respect of the partition

$$\beta_1 \dots \beta_k \alpha_{k+1} \dots \alpha_n / \alpha_1 \dots \alpha_k \beta_{k+1} \dots \beta_n,$$

which is obtained from $\alpha_1\alpha_2 \dots \alpha_m / \beta_1\beta_2 \dots \beta_n$ by the interchange of the groups $(\alpha_1\alpha_2 \dots \alpha_k)$ and $(\beta_1\beta_2 \dots \beta_k)$. The required derivate is

$$\phi(x, y) = (\alpha_{k+1} x) \dots (\alpha_m x)(\beta_{k+1} y) \dots (\beta_n y)\, \psi(x, y),$$

where

$$(xy)\, \psi(x, y) = (\alpha_1 x) \dots (\alpha_k x)(\beta_1 y) \dots (\beta_k y) - (\alpha_1 y) \dots (\alpha_k y)(\beta_1 x) \dots (\beta_k x).$$

The form $\psi(x, y)$ is the symmetric $(k-1, k-1)$ form, in respect of which any two of the points $\alpha_1\alpha_2 \dots \alpha_k$, and also any two of $\beta_1\beta_2 \dots \beta_k$ are correspondents. If $(\nu_1\nu_2 \dots \nu_k)$ is any group of the involution determined by $(\alpha_1\alpha_2 \dots \alpha_k)$ and $(\beta_1\beta_2 \dots \beta_k)$, then it is easy to see that any two of the ν's are also correspondents in ψ. Incidentally we see that the fixed points of ψ are the roots of the Jacobian of the involution. Thus

1·17 The derivate of the form $(\alpha_1 x) \dots (\alpha_m x)(\beta_1 y) \dots (\beta_n y)$ in respect of the partition $\beta_1 \dots \beta_k \alpha_{k+1} \dots \alpha_m / \alpha_1 \dots \alpha_k \beta_{k+1} \dots \beta_n$ is

$$(\alpha_{k+1} x) \dots (\alpha_m x)(\beta_{k+1} y) \dots (\beta_n y)\, \psi(x, y),$$

where ψ is the unique symmetric $(k-1, k-1)$ form, in respect of which any two of $\alpha_1\alpha_2 \dots \alpha_k$ and any two of $\beta_1\beta_2 \dots \beta_k$ are correspondents.

It follows that if the derivate of F in respect of a particular partition be ϕ, its derivate in respect of any other partition differs from ϕ by an expression of the type obtained in 1·17.

§ 2. Complete and Closed Sets of a Correspondence.

(6) *Canonical Forms with Complete Sets.*

A set of r points $i_1 i_2 \dots i_r$ will be said to form a *complete set* of an (m, n) correspondence F if every pair of points i, whether distinct or identical, are correspondents in F. The necessary and sufficient conditions for this are therefore the r^2 equations

$$F(i_p, i_q) = 0; \quad (p, q = 1, 2, \dots r).$$

The set $i_1 i_2 \dots i_r$ may be said to be a "semi-complete set" of F if every pair of *different* points i are correspondents in F; that is, if the $r(r-1)$ equations $F(i_p, i_q) = 0$ are satisfied for

$$p, q = 1, 2, \dots r; \; p \neq q.$$

Digitized by Google

The same set of r points $i_1 i_2 \dots i_r$ will be said to be a *closed set* of F if the m roots of each of the r equations $F(x, i_p) = 0$, and the n roots of each of the r equations $F(i_p, y) = 0$ are all included among $i_1 i_2 \dots i_r$; that is, if no point which is not an i has a point i for a correspondent in F. In other words, the set $i_1 i_2 \dots i_r$ is closed in regard to F, if starting from any i it is impossible to go out of the set by means of F.

It follows from the first definition that a complete set is necessarily a subset of the set of fixed points. This need not be true of a closed set. A closed set when it exists need not be complete, nor a complete set closed. There may however exist sets which are both complete and closed. As an obvious instance, an n-set complete in regard to an (n, n) correspondence is also closed with respect to it.

An (m, n) form F possessing a complete set $(i_1 i_2 \dots i_r)$ can be expressed in the shape:

2·1

$$F(x, y) = (i_1 x) \dots (i_r x)[P_1(x^{m-r}, y^n)] + (i_1 y) \dots (i_r y)[P_2(x^m, y^{n-r})].$$

For the right-hand side vanishes when $x = i_p$, $y = i_q$, and therefore possesses the complete set $i_1 i_2 \dots i_r$. Now the number of available parameters in F is

$$(m + 1)(n + 1) - r^2,$$

while the number of parameters on the right is

$$(m + 1)(n + 1) - r^2 + (m + 1 - r)(n + 1 - r).$$

This excess number $(m + 1 - r)(n + 1 - r)$ may be accounted for by the existence of identical relations, involving

$$(m + 1 - r)(n + 1 - r)$$

parameters, of the form

$$(i_1 x) \dots (i_r x)[(i_1 y) \dots (i_r y)\, Q(x^{m-r}, y^{n-r})] - (i_1 y) \dots (i_r y)[(i_1 x) \dots (i_r x)\, Q(x, y)] \equiv 0.$$

Thus the right side of 2·1 involves effectively the same number of parameters as F, and is consequently adequate for representing it.

The representation 2·1 can be rendered unique by imposing either the condition that P_1 should be identically apolar to $(i_1 y) \dots (i_r y)$, or the condition that P_2 should be identically apolar to $(i_1 x) \dots (i_r x)$.

(7) *Reduction of Rank through a Complete Set.*

2·2 Supposing for convenience that $m \geqslant n$, we go on to shew that if an (m, n) form F $(m \geqslant n)$ possesses a complete r-set, where of course $r \not> n$, and if $r > \dfrac{m+1}{2}$, $r > m - n$, then F suffers a reduction in rank from $n + 1$ to $(n + 1) - (2r - \overline{m+1})$.

Digitized by Google

For, the first term in the representation 2·1 being the product of a binary form and a $(m-r, n)$ form, its rank is the smaller of $m-r+1$, $n+1$. Similarly the rank of the second term is the smaller of $n-r+1$, $m+1$, that is, since $m \geqslant n$, it is $n-r+1$. Hence the rank of F cannot exceed

$$(n-r+1) + \text{smaller of } (m-r+1, n+1).$$

If n is smaller than $m-r$ this statement is trivial. Supposing then that $m-r<n$ or $r>m-n$, the rank of F cannot exceed

$$(n+1) - (2r-\overline{m+1}).$$

This statement again will be trivial unless $2r>m+1$. Thus, if $r>m-n$ and $2r>m+1$, there is the actual reduction stated in the rank of F.

Putting $r=n$ in 2·2, we have the special case,

2·3 An (m, n) form F possessing a complete n-set is of rank $m-n+2$, if $2n-1>m \geqslant n$.

A still more special case which has application to the pedal correspondence in n dimensions is:

2·4 An $(n+1, n)$ form possessing a complete n-set is of rank three $(n>1)$.

(8) *Geometrical Interpretation in n Dimensions.*

These theorems about the reduction of rank may be translated into simple geometrical properties of curves in n-dimensional space. The (m, n) form $F(x, y)$ $(m \geqslant n)$ institutes an (m, n) correspondence F between the points of a rational norm curve R_n in a space S_n of n dimensions. To any point x of R_n F makes correspond n points $y_1 y_2 \dots y_n$, and therefore the unique prime $y_1 y_2 \dots y_n$. As x traces out R_n, the corresponding prime $y_1 y_2 \dots y_n$ osculates a curve $\mathbf{R}_m$ of class m, which we may call the *associated envelope* of F; this associated envelope is evidently in birational correspondence with the norm curve R_n, and is therefore a rational envelope.

If F possesses a complete n-set $i_1 i_2 \dots i_n$ then, in this birational correspondence between R_n and $\mathbf{R}_m$, the prime $i_1 i_2 \dots i_n$ corresponds to each of the points $i_1, i_2 \dots i_n$. Hence it is a multiple prime, of multiplicity n, of the associated envelope. On reciprocating $\mathbf{R}_m$ we obtain a curve of order m, possessing an n-ple point, in S_n. We see therefore that 2·3 asserts that:

2·5 A curve of order m in S_n, possessing an n-ple point, must be completely contained in an S_{m-n+1} $(m \leqslant 2n-1)$.

This is geometrically obvious, since the S_{m-n+1} which joins the n-ple point to any other $m-n+1$ points of the curve cuts it in $m+1$ points, and therefore contains it completely. Similarly it is seen that 2·2 asserts that:

Digitized by Google

2·6 If a curve of order m ($\geqslant n$) in S_n, possesses a space L_{n-r} of $n-r$ dimensions, cutting it in r points, and if

$$r > m - n, \quad 2r > m + 1,$$

the curve must be completely contained in an $S_{m+n+1-2r}$.

The proof is exactly similar; the $S_{m+n+1-2r}$ which joins the L_{n-r} to any other $m+1-r$ points of the curve, cuts it in $m+1$ points, and therefore contains it completely.

(9) $(n+1, n)$ *Forms possessing a Complete and Closed n-Set.*

Let the $(n+1, n)$ form $F(x, y)$ possess the complete and closed n-set given by $f(x) = (i_1 x)(i_2 x) \ldots (i_n x)$. By 2·1, the form F can be written in the shape:

$$F(x, y) = f(x) \,.\, (ax)(by)^n + f(y) \,.\, (cx)^{n+1},$$

where (a, b) and c are symbolic. If we now use the condition that $f(x)$ is also a closed set, we see that the n points $y = i_1, i_2 \ldots i_n$ correspond in $(ax)(by)^n$ to the n points $x = i_1, i_2 \ldots i_n$ in some order. Thus the $(1, n)$ correspondence $(ax)(by)^n$ merely effects a substitution ω on the n points $y = i_1, i_2 \ldots i_n$. Hence, *there are* $n!$ *ways in which a complete n-set can be closed in regard to an* $(n+1, n)$ *correspondence, each of these ways corresponding to a permutation* ω *on the n points of the set.*

Two types of closure appear to be noteworthy; the first is that in which ω is the identical permutation, the second is that in which n is even, and ω interchanges the n points in pairs. It is the latter type which, when $n = 2$, occurs in the case of the pedal form.

Digitized by Google

The stability of a strut under thrust, when buckling is resisted by a force proportional to the displacement. By S. GOLDSTEIN, B.A., St John's College.

[*Received* 2 February, *read* 1 March 1926.]

1. The problem discussed in this paper is that of the stability of a strut of length l, either clamped at both ends or pinned at both ends, under a thrust R, with buckling resisted by a force equal to K times the displacement. The problem was suggested to me by Mr H. Jeffreys, of St John's College, as being of some geophysical interest (see 6). It later appeared that it might have some connection with recent work on the collapse of thin films on water (see 9).

2. *Fundamental Differential Equation.*

The usual elementary theory is used*. The strut being of fixed length, we know by the principle of the exchange of stabilities that when R is gradually increased from zero till a position of equilibrium of the strut other than the straight form just becomes possible, then the straight form itself becomes unstable. We therefore seek for the relation between R and l which holds when such a position of equilibrium exists. The axis of x is taken along the central line of the strut in its unbuckled state, which is also taken to be the line of thrust. The axis of y is in the plane of buckling.

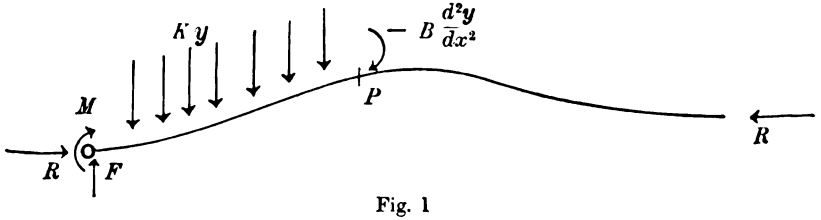


Fig. 1

Consider first a clamped strut. Take the origin O at one end, and let P be a point whose abscissa is x, the strut being supposed just buckled. Take moments about P for the portion OP of the strut. This gives the equation

$$B\frac{d^2y}{dx^2} + Ry + \int_0^x Ky(\xi)(x-\xi)\,d\xi - M - Fx = 0 \quad \ldots(1),$$

where M is the bending moment and F the shearing force at the clamp, as shown (Fig. 1): B is the flexural rigidity of the strut, and $y(\xi)$ is the value of y at a distance ξ from the end.

* Prescott, *Applied Elasticity*, pp. 82 *et seq.* (1924).

Digitized by Google

Differentiating this equation twice with respect to x gives

$$B\frac{d^4y}{dx^4} + R\frac{d^2y}{dx^2} + Ky = 0 \cdot \qquad \ldots\ldots(2).$$

This equation is unaltered if the origin be taken at any other point along the central line; and the same equation clearly holds for a pinned rod.

3. *Clamped Strut:* $R^2 < 4KB$.

If $R^2 < 4KB$ the equation

$$Bm^4 + Rm^2 + K = 0 \qquad \ldots\ldots(3)$$

has roots of the form $\pm\alpha \pm i\beta$, where α and β are real and positive.

Let the ends be $x = 0$ and $x = l$. At the ends y and dy/dx vanish. The solution of (2) is

$$y = A\cosh\alpha x\cos\beta x + B\cosh\alpha x\sin\beta x + C\sinh\alpha x\cos\beta x + D\sinh\alpha x\sin\beta x \ldots\ldots(4).$$

By the conditions at $x = 0$, $A = 0$ and $\beta B + \alpha C = 0$, and we can write

$$y = D\sinh\alpha x\sin\beta x - \alpha H\cosh\alpha x\sin\beta x + \beta H\sinh\alpha x\cos\beta x \qquad \ldots\ldots(5).$$

The conditions at $x = l$ give, on eliminating H/D and simplifying,

$$\beta^2\sinh^2\alpha l - \alpha^2\sin^2\beta l = 0,$$

or

$$\beta\sinh\alpha l = \pm\alpha\sin\beta l \qquad \ldots\ldots(6).$$

This cannot be satisfied with $l > 0$, since $\sinh\alpha l > \alpha l$ and $\sin\beta l < \beta l$.

Hence, *for values of R less than $\sqrt{4KB}$, there is stability for any length.*

4. *Clamped Strut:* $R^2 > 4KB$.

For $R^2 > 4KB$, take the origin in the middle. Then y and dy/dx vanish at $x = \pm\frac{1}{2}l$.

The roots of (3) are $\pm i\gamma$, $\pm i\delta$, where

$$B\gamma^2 = \tfrac{1}{2}R + \sqrt{(\tfrac{1}{4}R^2 - KB)} \qquad \ldots\ldots(7),$$

$$B\delta^2 = \tfrac{1}{2}R - \sqrt{(\tfrac{1}{4}R^2 - KB)} \qquad \ldots\ldots(8),$$

γ and δ being real and positive.

The solution of (2) is then

$$y = A\cos\gamma x + B\sin\gamma x + C\cos\delta x + D\sin\delta x \quad \ldots(9),$$

and the end conditions give

$$\begin{aligned} A\cos\tfrac{1}{2}\gamma l + C\cos\tfrac{1}{2}\delta l &= 0, \\ B\sin\tfrac{1}{2}\gamma l + D\sin\tfrac{1}{2}\delta l &= 0, \\ \gamma A\sin\tfrac{1}{2}\gamma l + \delta C\sin\tfrac{1}{2}\delta l &= 0, \\ \gamma B\cos\tfrac{1}{2}\gamma l + \delta D\cos\tfrac{1}{2}\delta l &= 0. \end{aligned}$$

Digitized by Google

Hence either $B = D = 0, \quad \beta \sin \alpha l + \alpha \sin \beta l = 0$ .....(10),

or $A = C = 0, \quad \beta \sin \alpha l - \alpha \sin \beta l = 0$ .....(11),

where $\alpha = \frac{1}{2}(\gamma + \delta), \quad \beta = \frac{1}{2}(\gamma - \delta)$ .....(12).

The expressions for y in the two cases are of the forms

$$\left.\begin{aligned} y &= E(\cos \tfrac{1}{2}\delta l \cos \gamma x - \cos \tfrac{1}{2}\gamma l \cos \delta x) \text{ corresponding to (10)} \\ y &= F(\sin \tfrac{1}{2}\delta l \sin \gamma x - \sin \tfrac{1}{2}\gamma l \sin \delta x) \text{ corresponding to (11)} \end{aligned}\right\} \quad (13).$$

To consider the equation $\beta \sin \alpha l = \pm \alpha \sin \beta l$ $(\alpha > \beta)$ let

$$\beta l = x, \quad \beta/\alpha = p \quad (p < 1) \quad \text{.....(14).}$$

The equation is $\pm p \sin (x/p) = \sin x$ .....(15).

We are concerned only with positive values of x.

Let $f(x) = \sin x - p \sin (x/p)$. Then, if r is an odd integer, we have

$$f(r\pi - \sin^{-1} p) > 0, \quad f(r\pi + \sin^{-1} p) < 0,$$

and, if r is an even integer,

$$f(r\pi - \sin^{-1} p) < 0, \quad f(r\pi - \sin^{-1} p) > 0.$$

Hence there is a root of $f(x) = 0$ in each of the intervals

$$(r\pi - \sin^{-1} p, \quad r\pi + \sin^{-1} p).$$

Also $f(0) = 0$, $f'(x) > 0$ for $x < \frac{1}{2}\pi$. Therefore $f(x) > 0$ for $x < \frac{1}{2}\pi$. For $\frac{1}{2}\pi < x < \pi - \sin^{-1} p$, $\sin x > p$, and therefore $f(x) > 0$. Hence $f(x) > 0$ for $0 < x < \pi - \sin^{-1} p$.

For $\pi + \sin^{-1} p < x < 2\pi - \sin^{-1} p$, $\sin x < -p$ and $f(x) < 0$; for $2\pi + \sin^{-1} p < x < 3\pi - \sin^{-1} p$, $\sin x > p$ and $f(x) > 0$; and so on. Hence there is no root of $f(x) = 0$ outside the intervals

$$(r\pi - \sin^{-1} p, \quad r\pi + \sin^{-1} p).$$

Similarly there is a root of $\sin x + p \sin x/p = 0$ in each of the intervals $(r\pi - \sin^{-1} p,\ r\pi + \sin^{-1} p)$, and no root outside these intervals. There is only one root of each equation in each interval, since otherwise we should have $\frac{1}{2}p\pi < \sin^{-1} p$. The root of one equation is in the first half of each interval, the other equation having its root in the second half. The first equation

$$(p \sin x/p = \sin x)$$

has its first root the smaller if the greatest integer contained in p^{-1} is even: it has its second root the smaller if the greatest integer contained in $2p^{-1}$ is odd: its third root the smaller if the greatest integer in $3p^{-1}$ is even, and so on.

All these results are easily obtained by drawing the graphs of $\sin x$ and $\pm p \sin x/p$ for a few values of p.

The first root for x is then certainly between $\frac{1}{2}\pi$ and π, so that βl lies between $\frac{1}{2}\pi$ and π. For the second root βl is between π and $\frac{3}{2}\pi$: for the third root between $\frac{3}{2}\pi$ and 2π, and so on. Now β

Digitized by Google

vanishes when $R^2 = 4KB$ and increases with R, so that if we suppose l to remain fixed and R to increase from $2\sqrt{KB}$, instability will first manifest itself when the first root is reached, for which βl lies between $\frac{1}{2}\pi$ and π. Further, if l is comparable with continental dimensions the first root will be reached by βl for a value of R very near to $2\sqrt{KB}$ (see 7).

5. *Clamped Strut: wave-length and number of nodes after buckling.*

The expressions for y in (13) can also be written in the forms

$$\left.\begin{aligned} y/E &= \sin\beta(\tfrac{1}{2}l+x)\sin\alpha(\tfrac{1}{2}l-x) + \sin\beta(\tfrac{1}{2}l-x)\sin\alpha(\tfrac{1}{2}l+x) \\ y/F &= \sin\beta(\tfrac{1}{2}l+x)\sin\alpha(\tfrac{1}{2}l-x) - \sin\beta(\tfrac{1}{2}l-x)\sin\alpha(\tfrac{1}{2}l+x) \end{aligned}\right\} \quad (16),$$

and we may get some notion of the number of nodes corresponding to the various roots of (15) by taking $\beta l = m\pi$, where m is a positive integer. This reduces the forms (13) or (16) to

$$\left.\begin{aligned} y &= G\cos m\pi x/l \cos m\pi x/pl \ (m \text{ odd}) \\ &= -G\sin m\pi x/l \sin m\pi x/pl \ (m \text{ even}) \end{aligned}\right\} \text{ for the symmetrical modes} \quad \ldots\ldots(17),$$

$$\left.\begin{aligned} y &= H\cos m\pi x/l \sin m\pi x/pl \ (m \text{ odd}) \\ &= -H\sin m\pi x/l \cos m\pi x/pl \ (m \text{ even}) \end{aligned}\right\} \text{ for the asymmetrical modes} \quad \ldots\ldots(18).$$

For $m = 1$ the number of internal nodes is of the order of p^{-1} and the wave-length is of the order of $2pl$. For higher values of m the number of nodes is usually somewhat greater than mp^{-1}. We may compare this with the results for $K = 0$, in which case there are no nodes when buckling first occurs, one node for the first symmetrical mode, two for the second asymmetrical mode, and so on. (Of course, the higher modes have practical application only if the lower ones are prevented by constraints.)

6. *Note on the Formation of Mountains by Horizontal Compression.*

According, for example, to the thermal contraction theory of mountain formation, the process of the cooling of the earth was such as to leave the crust in a state of horizontal compression, under which it gave way by crumpling. The objection has been raised* that horizontal compression would give deformations only on a very small scale, a thin crust being unable to transmit considerable stresses across regions of continental extent without buckling. We shall see that for buckling to occur the depth would have to be a good deal less than is generally considered admissible. The depth with which we are concerned is the depth of the crust down to the level of no strain. The criticism applies not only to

* E.g. Bowie, *J. Franklin Inst*, vol. 198, p. 188 (Aug. 1924).

Digitized by Google

the thermal contraction theory, but to any theory of horizontal compression.

In buckling the crust would have to take with it the matter underneath, so that the problem is substantially the same as the one that we have been considering. Further, for regions of continental extent, the breadth is sufficiently large compared with the depth to prevent buckling horizontally, so that our two-dimensional theory applies.

To obtain an estimate of the depth required, we take a strut of rectangular section of depth d and breadth b. Then $B = EI$, where E is Young's modulus and I is the moment of inertia of a cross-section about the neutral line in its plane; $I = \frac{1}{12}bd^3$ and $E = 6 \times 10^{11}$ dynes/cm.^{2} roughly, so that $B = 5 \times 10^{10} bd^3$ dynes $\times$ cm.^{2} $K = bg\rho$, where ρ is the density of the material which the crust must move in buckling. Then $K = 3b \times 10^3$ dynes/cm.^{2} roughly, and $4BK = 6 \times 10^{14} b^2 d^3$ dynes^2. In order that buckling may be produced before crumpling, we must have both R/bd less than the breaking stress, which is nearly 19 dynes/cm.^{2}, and R greater than $2\sqrt{BK}$. Hence we must have

$$6 \times 10^{14} b^2 d^3 < R^2 < 10^{18} b^2 d^2, \text{ or } d < 1{\cdot}6 \times 10^3 \text{ cm.};$$

i.e. *the depth of the level of no strain must not exceed* 16 *metres*! The present depth of the level of no strain is between 75 and 80 kilometres, and its depth at the last epoch of mountain formation was not much different*.

It follows that the stabilizing effect of gravity is such as to ensure that any possible horizontal stress can be transmitted without buckling, and yield will therefore occur first by fracture at the weakest point; this is what is actually needed for mountain formation.

7. *Numerical Example of the Buckling of a Clamped Strut.*

For America l is of the order of 4×10^8 cm., and for the Alpine fold of order 3×10^7 cm. Hence, if we take the depth to be less than the critical depth, the lowest value of R that would produce buckling would correspond to a value of β of order, say 10^{-7} cm.$^{-1}$, and would be very near to $2\sqrt{KB}$. If we take $d = 10^3$ cm. we have $B = 5b \times 10^{19}$ dynes $\times$ cm.^{2}, $4BK = 6b^2 \times 10^{23}$ dynes^2 and the critical value of R is $7{\cdot}75b \times 10^{11}$ dynes. If we take $R = 7{\cdot}8b \times 10^{11}$ dynes we get

$$\gamma = 9{\cdot}3 \times 10^{-3} \text{ cm.}^{-1}, \quad \delta = 8{\cdot}3 \times 10^{-3} \text{ cm.}^{-1},$$

$$\alpha = 8{\cdot}8 \times 10^{-3} \text{ cm.}^{-1}, \quad \beta = 5 \times 10^{-4} \text{ cm.}^{-1}, \quad p = 0{\cdot}06$$

and $\sin^{-1} p = 0{\cdot}06$. For l of the order of 10^7 cm. the value of β, and therefore also of R, is already too large. This shows how near

* Jeffreys, *Phil. Mag.* vol. 32, pp. 583 and 585 (Dec. 1916).

Digitized by Google

R must be to its critical value $2\sqrt{BK}$ for l of anything like this order of magnitude, and how near βl must be to π for the first root in such cases. (β and p both continually increase from zero as R increases from $2\sqrt{BK}$.)

The value of R chosen will be just large enough to cause buckling for a length of the order of $2\pi \times 10^5$ cm. The wave-length would then be of the order of $7{\cdot}54 \times 10^3$ cm., and the number of internal nodes would be 16 or 17.

We shall see in 9 that the results for a pinned strut are practically the same.

(NOTE.—If the wave-length does not come out to be larger than, or at any rate comparable with, the depth, the theory certainly ceases to have any practical significance.)

8. *Strut pinned at both ends:* $R^2 < 4KB$.

Equation (2) still holds. Take the origin in the middle. The end conditions are $y = 0$, $d^2y/dx^2 = 0$ at $x = \pm \frac{1}{2}l$.

If $R^2 < 4KB$ we may write, as before,

$$y = A \cosh \alpha x \cos \beta x + B \cosh \alpha x \sin \beta x + C \sinh \alpha x \cos \beta x + D \sinh \alpha x \sin \beta x \ (\alpha, \beta \text{ real and positive}).$$

The end conditions give

$$\left.\begin{array}{l} \text{either } B = C = 0, \ \cosh^2 \frac{1}{2}\alpha l \cos^2 \frac{1}{2}\beta l + \sinh^2 \frac{1}{2}\alpha l \sin^2 \frac{1}{2}\beta l = 0 \\ \text{or} \quad A = D = 0, \ \cosh^2 \frac{1}{2}\alpha l \sin^2 \frac{1}{2}\beta l + \sinh^2 \frac{1}{2}\alpha l \cos^2 \frac{1}{2}\beta l = 0 \end{array}\right\} \quad (19),$$

and again there cannot be buckling if $R^2 < 4KB$.

9. *Pinned Strut:* $R^2 > 4KB$.

If $R^2 > 4KB$ we may write

$$y = A \cos \gamma x + B \sin \gamma x + C \cos \delta x + D \sin \delta x,$$

where γ and δ are given by (7) and (8).

The end conditions give

$$\begin{aligned} A \cos \tfrac{1}{2}\gamma l + C \cos \tfrac{1}{2}\delta l &= 0, \\ B \sin \tfrac{1}{2}\gamma l + D \sin \tfrac{1}{2}\delta l &= 0, \\ \gamma^2 A \cos \tfrac{1}{2}\gamma l + \delta^2 C \cos \tfrac{1}{2}\delta l &= 0, \\ \gamma^2 B \sin \tfrac{1}{2}\gamma l + \delta^2 D \sin \tfrac{1}{2}\delta l &= 0. \end{aligned}$$

Since, in general, $\gamma \neq \delta$, it appears that we must have

$$A \cos \tfrac{1}{2}\gamma l = B \sin \tfrac{1}{2}\gamma l = C \cos \tfrac{1}{2}\delta l = D \sin \tfrac{1}{2}\delta l = 0 \quad (20).$$

Hence three of A, B, C, D must be zero and either γl or δl must be equal to $n\pi$, where n is a positive integer.

As R increases from $2\sqrt{KB}$, γ increases and δ decreases, and buckling will first take place when γ or δ becomes equal to a multiple of π/l. We have then that

$$y = A \cos n\pi x/l \text{ or } B \sin n\pi x/l \quad \ldots\ldots\ldots(21),$$

Digitized by Google

where $n\pi/l$ is the multiple of π/l first reached by

$$\{\tfrac{1}{2}R/B \pm \sqrt{(\tfrac{1}{4}R^2/B^2 - K/B)}\}^{\frac{1}{2}},$$

as R increases from $2\sqrt{KB}$. The wave-length is $2l/n$ and the number of internal nodes is $n - 1$.

When l is of the order $2\pi \times 10^5$ cm. and d is 10^3 cm., so that $4BK = 6b^2 \times 10^{23}$ dynes^2 and R increases from $7{\cdot}75b \times 10^{11}$ dynes as before (see 7), γ increases and δ decreases from $8{\cdot}8 \times 10^{-5}$ cm.$^{-1}$, while π/l is 5×10^{-6} cm.$^{-1}$ When $R = 7{\cdot}8b \times 10^{11}$ dynes, $\gamma = 9{\cdot}3 \times 10^{-4}$ cm.$^{-1}$ and $\delta = 8{\cdot}3 \times 10^{-4}$ cm.$^{-1}$, so that a multiple of π/l is reached for a value of R nearer the critical value. Thus the pinned strut buckles before a clamped strut of the same length would do so. Also, for the case considered, n is 17 or 18, so that the wave-length and the number of nodes are about the same as for a clamped strut.

10. *The Compression and Collapse of Surface Films on Water.*

In a series of papers* Mr N. K. Adam has discussed the behaviour, under compression, of surface films on water. These films are of one molecule thickness, and when the area per molecule is sufficiently reduced (to about $2{\cdot}2 \times 10^{-15}$ cm.^{2}) the film behaves like a solid and there is a linear relation between the compression and the thrust. The linear relation holds when the thrust is increased from 3 or 4 dynes/cm. until we reach the stage at which the film gives way†.

The atoms being in equilibrium under their mutual attractions and repulsions, we suppose that when the system is compressed, the restoring force on any atom is approximately proportional to its relative displacement, and the usual elementary elastic theory may be applied. First, we can calculate from the compressibility the value of E, Young's modulus for the film. Then if the film were to buckle, it would have to raise and lower the water with it, so that we can employ the theory of 2–5, 8 and 9 to find under what thrust instability should first manifest itself. The first effect would be to cause buckling, which in a film of one molecule thickness would necessarily be imperceptible; but very shortly after this the film should give way. We assume that the film does not collapse before the thrust reaches its theoretical value for buckling.

11. *The Value of Young's Modulus for a Solid Film.*

We deal with a breadth of 1 cm. Then‡, since any error in experiment will increase the observed compressibility, we shall use the lowest value obtained, 0·4 per cent. for a thrust of 10 dynes.

* *Proc. Roy. Soc.* A, vol. 99, p. 336 (1921); vol. 101, p. 452 (1922), etc. References to the two papers mentioned will be given as Part I and Part II.

† Part I.

‡ Part II, p. 458.

Digitized by Google

We may take the depth of a film as being of the order of $2{\cdot}5 \times 10^{-7}$ cm.* Then E is equal to the force per unit depth divided by the compressibility. This gives $E = O(10^{10})$ dynes/cm.[2]

12. *The Collapse of a Solid Film.*

The value of the thrust that produces collapse cannot be determined with accuracy, but is anything from 20 to 45 dynes†. If we suppose buckling to precede collapse the value of R that gives buckling should not be very different from the value for collapse, preference being given to the smaller values.

Now $B = EI$, $I = \frac{1}{12}d^3$, $K = g$, and with our data and the above value for E, $B = O(10^{-11})$ dynes $\times$ cm.[2], $BK = O(10^{-8})$ dynes[2].

If we were to take R to be 16 dynes, we should, employing the results of 4, find that $\beta = O(4{\cdot}5 \times 10^{5})$ cm.$^{-1}$ This value of β is much too large for a length anywhere between 2 or 3 cm. to 20 or 30 cm. We want a value of β not greater than about 1 cm.$^{-1}$, and this means a value of R of the order of 10^{-4} dynes, which is nonsense. With R 16 dynes, values of B between 6×10^{-2} and $6{\cdot}4 \times 10^{-2}$ dynes $\times$ cm.[2] give values of β between 2 and 0, so that for this value of R we should want $B = O(6 \times 10^{-2})$ dynes cm.[2], $E = O(6 \times 10^{19})$ dynes/cm.[2]

The results if we considered the film as pinned would be practically the same, and would show the same discrepancy.

Clearly then there are other factors to be considered that make for stability.

13. *The Effect of Surface Tension.*

Since in both compression and buckling the area of the water-film surface is changed, we may expect surface tension to play a part. For compression, if T is the surface tension,

$$E = (R + T)/(d \times \text{compressibility}).$$

Since the surface tension of a solid film cannot vary appreciably with additional compression, we may‡ take T to be 70 dynes/cm. at temperatures between 10° and 20° C., the temperature of Mr Adam's experiments on palmitic acid§. Then

$$E = O(8 \times 10^{10}) \text{ dynes/cm.}^2$$

The equation (2) becomes

$$B\frac{d^4y}{dx^4} + (R - T)\frac{d^2y}{dx^2} + Ky = 0 \qquad \text{.....(22).}$$

The value of R that produces buckling must exceed T. According to the experimental results, however, the film collapses for values of R (20–45 dynes/cm.) less than T, and therefore

* Part II, p. 463. † Part I, p. 340.
‡ Cary and Rideal, *Proc. Roy. Soc.* A, vol. 109, p. 320 (1925).
§ Part I, p. 340.

Digitized by Google

less than the theoretical value for buckling. Hence instability must first manifest itself by a weakness in the film causing collapse; this happens before the film has a chance to buckle.

14. *The Frequency of Vibration of an Atom in a Solid Film.*

From the value of E we can make a very rough calculation for the frequency of vibration of an atom in a solid film. The cross-section of the head of a molecule is of the order of 26×10^{-16} cm.^{2}*, i.e. $(5 \times 10^{-8}$ cm.$)^2$. We take 5×10^{-8} cm. as the original distance D, between two atoms, that is decreased to $D - \delta$ by compression. Then we have

$$\text{Force per unit area across a cross-section} = E\delta/D.$$

If n is the number of atoms in the molecule, d the depth of the unimolecular layer, then the restoring force on an atom

$$= (E\delta/D) \div (n/dD) = Ed\delta/n.$$

The frequency is $\sqrt{\mu}/2\pi$, where

$$\mu = \text{restoring force} \div (\delta \times \text{mass}).$$

The mass of an atom is $1{\cdot}65 \times 10^{-24} A$ gm., where A is its atomic weight. Take $A = 12$ (for a carbon atom), and $n = 60$, for example. With $E = O(8 \times 10^{10})$ dynes/cm.^{2}, $\mu = O(1{\cdot}7 \times 10^{25})$ and the frequency $= O(6{\cdot}5 \times 10^{11})$.

We may compare this with the corresponding frequency for crystals, as obtained by the method of residual rays†. The frequencies are c/λ, where c is the velocity of light $(= 3 \times 10^{10}$ cm./sec.) and λ, the wave-length of the residual rays, is between $1{\cdot}2 \times 10^{-3}$ and $1{\cdot}2 \times 10^{-2}$ cm. for most crystals. This gives frequencies varying between $2{\cdot}5 \times 10^{13}$ and $2{\cdot}5 \times 10^{12}$.

15. *Summary.*

For a strut under thrust when buckling is resisted by a force proportional to the displacement, it is shown that for both a clamped and a pinned strut there is stability for any length when the thrust is less than a certain critical value; and the relation is found between the length and the first value of the thrust above this that will give buckling. A notion of the wave-length and number of nodes after buckling is obtained.

The results are applied to the theory of the formation of mountains by horizontal compression, and it is shown that the crust of the earth would be able to transmit, without buckling, stresses right up to the breaking stress across regions of continental extent, unless the depth down to the level of no strain were less than 16 metres.

* Part II, p. 463.

† Born, *Dynamik der Kristallgitter*, p. 623 (1923).

Digitized by Google

An attempt is made to employ the results also to consider the stability of solid surface films under compression, and it is seen that other factors, making for stability, must enter besides rigidity and gravity. This is supplied by surface tension, and it then appears that collapse due to a weakness in the film must precede buckling.

A rough calculation is given of the frequency of vibration of an atom in a solid film.

16. I have to thank Mr H. Jeffreys for advice and assistance in the preparation of this paper, and Mr E. K. Rideal for showing me an experiment similar to those of Mr Adam.

Digitized by Google

Insular Gravity and Oceanic Isostasy. By Sir JOSEPH LARMOR, Lucasian Professor.

[*Received* 8 February 1926.]

A century ago geodetic and gravitational universal surveys were mainly concerned with determining the effective (gravitational) ellipticity of the Earth, after due allowance had been made for local anomalies, with especial view to the exact purposes of physical astronomy. One of the chief of these anomalies was exhibited by a remark of Airy, after scrutiny of the available data in his treatise (1830) on "Figure of the Earth" in the *Encyclopedia Metropolitana*, that the observations show gravity to be abnormally in excess on island stations. It appeared, for instance, that this cause might make the mass of the Moon uncertain up to 2 per cent. A very refined explanation of this anomaly of island stations (which will be seen presently to be only partially effective) was offered by Sir George Stokes, from whom this last remark is quoted, in the course of a memoir*, fundamental for theoretical geodesy, in which he demonstrated that no outside survey could lead to any certain knowledge of the distribution of mass inside the Earth, even in its outer crust, except as a matter of probability when backed up by geological knowledge.

It is explained there that the form of the sea-level must be locally depressed over a deep ocean, owing to defect of density; and in consequence on insular stations gravity at sea-level is measured abnormally nearer to the centre of the Earth as a whole, so that from this cause its value is greater than that belonging to the mean spheroidal surface. In fact, the form of the ocean is an equipotential surface, including therein the potential of the centrifugal force of rotation in the familiar manner: but the part of the potential arising from the local water is abnormally small on account of its low density, and this defect must, in absence of local compensation, be made up by a greater potential of the Earth as a whole, which demands depression of the local ocean surface towards the Earth's centre.

The opposite result would arise from excess matter of an adjacent mountain or island peak: that would raise the ocean level in its vicinity and thereby indirectly diminish gravity, measured at sea-level, as determined by levelling operations.

For example, at the centre of a circular oceanic basin of radius

* *Cambridge Transactions* (1849): reprinted in *Math. and Phys. Papers*, vol. II. Some idea of the great debt owed by the Indian and other gravitational surveys to the continuous amateur advice of Sir G. G. Stokes, spread over half a century of their development, may be gleaned from the collection of his *Scientific Correspondence* (Camb. Univ. Press), vol. II, pp. 253–325.

Digitized by Google

b and uniform depth h, its defect of potential would be with sufficient accuracy $\int \gamma\rho' h 2\pi r\,dr/r$, where ρ' is the defect of density of the water below that of the average terrestrial crust; thus it is $2\pi\gamma\rho' bh$, where γ is the constant of gravitation given by $\gamma E/a^2 = g$. Here $E = \frac{4}{3}\pi a^3\rho$, ρ being $\frac{1}{2}$, is the mass of the Earth of radius a. As the potential of the Earth as a whole is $V = \gamma E/r$, this change of local potential, say δV_0, would be compensated by change of sea-level δh, where $\delta V_0/V = -\delta h/r$. Thus in the present case the fall of level relative to depth of ocean is given by the expression

$$-\frac{\delta h}{h} = \frac{a 2\pi\rho' b}{E/a} = \frac{3}{2}\frac{\rho'}{\rho}\frac{b}{a} = \frac{9}{22}\frac{b}{a},$$

while

$$\frac{\delta g}{g} = -2\frac{\delta h}{a}.$$

If the radius b of the oceanic basin is 50 miles this fall would be the fraction $\frac{9}{22}\cdot\frac{50}{4000}$ or $\frac{1}{200}$ of its depth; if the radius were larger it would increase in direct proportion until it is a considerable fraction of the Earth's radius. A cup-shaped ocean could be similarly treated.

The steady sea-level would thus be depressed by $\frac{1}{10}$ of a mile owing to local causes, at the centre of a basin of 500 miles radius and 2 miles deep, in free communication with the other oceanic waters: and this approach to the Earth's centre would involve increase of g measured at ocean level, given by $\delta g/g = -2\delta h/a$, or here $\delta g = \cdot 05$ cm./sec.^{2}, where g is about 981, which is over one-third of the order of magnitude of the observed excesses at island stations.

But this explanation fails because there is a predominant offset. The vertical attraction of the local ocean regarded as an extensive flat slab of water is abnormally small by $2\pi\gamma\rho' h$, where $g = \gamma E/a^2$, that is by $g\rho' 2\pi a^2 h/E$ or $\frac{3}{2}\frac{\rho'}{\rho}\frac{h}{a}g$; thus this direct defect in g may be much the greater, being $\frac{1}{2}a/b$ times the indirect excess. There is however some effect in the other direction due to excess density of the local land, which is usually a substantial correction. This preponderance destroys and even reverses the Stokes explanation of the oceanic anomaly. Indeed closer examination shows that, as based by him*, rather confusedly as it seems, it depends on a potential equation used by Laplace which can, in limited manner,

* *Math. and Phys. Papers*, vol. II, p. 153. Stokes did not make any correction in this reprint in 1883: but Dr Bowie states (*loc. cit. infra*) that there is no generally accepted explanation other than compensating excess of density beneath the ocean.

This analysis of Stokes in fact establishes as a general proposition that the effect of *distant* irregularities of surface mass consists of a direct vertical attraction, say g'', together with an indirect part due to change of level, equal to $-4g''$, thus countervailing four times: this influence, of wide range and presumably actually small, is superposed on the *local* effect here considered.

Digitized by Google

apply only to a locally infinitely thin spherical layer. The principle of depressed level became familiar, simple examples being worked out, *ab initio* and so correctly, by way of illustration in Chap. IV of Colonel A. R. Clarke's standard treatise on Geodesy (1880), from the point of view however only of levelling operations, not of gravity.

But soon the discussion of the data of the Indian geodetic survey, by Archdeacon Pratt in India, revealed new features*, by showing strong residual defect of gravity on the Himalayas, such as could only be accounted for by a large defect of density underneath the mountains. Airy's idea that the mountains might be buoyed up by extensive roots floating in a denser magma, existing beneath a *thin* crust, could not of course now be maintained, at any rate in that form, in view of the high rigidity of the Earth as a whole. But there was much to be said, on various counts, for a thinner and deeper viscid stratum, lying between the crustal material and the solid core, in which in the tendency towards equilibrium the pressure due to the weight of the crust must in course of ages have become equalized laterally, at any rate partially, and the load upon it thus made uniform to that degree everywhere. It is implied that there are no local abnormalities of density in the core, which is reasonable as the core is probably metallic. This is the hypothesis of isostasy, propounded as a universal principle by Dutton and worked out systematically by Hayford and his colleagues of the American Survey, who found that it gave a fair account of the usually slighter anomalies (mainly of levelling) revealed in that great undertaking†.

Circumspection is, however, suggested in applying these ideas to the anomalies at oceanic stations; for the Stokes explanation already claimed to be an effective *vera causa*, without aid from compensation of density underneath. It happens that the subject is amenable in a general way to simple elucidation: and as the essential circumstances for submarine mountains and landscapes can perhaps be more directly estimated, it seems indeed to provide in some respects a closer test. On an ideal very narrow island-peak of negligible mass, in a wide ocean of uniform depth, with adjustment as a whole to general isostasy by denser horizontal strata underneath, there would be but slight resultant abnormality of the local part of the attraction. For the totality of the strata could almost be regarded as an extensive thin flat sheet, while local defect of potential on which change of sea-level depends would be still more closely compensated by the extra mass below‡. Hence,

* In 1855–9: cf. A. R. Clarke, *Geodesy*, pp. 96–8.

† Cf. the chapter in H. Jeffreys' recent treatise *The Earth*.

‡ In the case illustrated above, with radius of ocean about 500 miles and depth of compensation 100 miles, about 10 per cent. of the anomaly both of attraction and of potential would remain after compensation of the ocean.

Digitized by Google

in contrast to the Stokes uncompensated case above, under isostatic conditions gravity and level ought both to be regular over a wide ocean of nearly uniform depth with strata nearly horizontal underneath.

Thus in considering gravity-data, say over the Pacific Ocean, we may on the isostatic hypothesis consider only this excess density, over that of water, of the local land distributions in an ocean with an ideal conveniently assigned flat bottom: and any gravity anomaly must be due to these excesses alone, together with isostatic compensations of opposite amount underneath to maintain the average total load for the crustal strata. The disturbance of ocean level due to them would be slight if they are merely local peaks. If the station is on a straight ledge even of a narrow island it will be considerable*.

This note, with its limitations as regards scope that are imposed by imperfect information, has been prompted by an important recent Report on "Isostasy in the Southern Pacific" by Dr W. Bowie† in which, from systematic calculation for the five insular stations that were examined, local isostasy is found to account for about three-fourths of the observed local excesses in g, which are there of the order of over 0·1 in 981. Since the Stokes effect of change of ocean level is found inadequate, the greater part of the excess of gravity ought to be due solely to direct attraction of the local excess of land density measured from an averaged flat ocean floor with a cancelling compensation underneath, if oceanic compensation is to hold good. A set of gravity determinations over the wide ocean, combined with soundings, would thus throw interesting light.

For a land survey, as in the Himalayas, heights determined by levelling operations are reckoned from an ideal ocean level which would be affected in the same way as the actual ocean level near island masses, that is only by local excess masses at the surface above a mean compensated distribution when the local compensations extend very deep.

There is a statement near the end of Dr Bowie's Report that, in the cases examined, if the densities of the island "pedestals" were 10 or 15 per cent. (say 12) above the normal value $2\frac{1}{2}$, about one-eighth of the remaining non-isostatic excess of gravity would be accounted for. From this we may perhaps infer that the excess density over that of water, namely $1\frac{1}{2}$, along with a cancelling diminution below on account of the now complete compensation, would account for $\frac{100}{12}.\frac{3}{5}.\frac{1}{8}$, or say $\frac{3}{5}$ of the whole local excess of gravity. This fraction is not discordant with the estimates (around

* Cf. A. R. Clarke, *Geodesy*, p. 93; or Thomson and Tait, *Nat. Phil.* (1867); also *infra* for conical forms—and the remark added at the end.

† *Proc. Washington Academy of Sciences* (Dec. 4, 1925).

Digitized by Google

$\frac{3}{4}$) in the Report, and is in so far confirmation. As the Stokes effect of change of ocean level is now itself obliterated by compensation, that over-compensation by about one-third could be accounted for, as Dr Bowie suggests, by an increase of mean excess density of crustal strata which need only be from $1\frac{1}{2}$ to about 2.

In illustration, for a conical island of height h and angle 2α the potential at its vertex is $\pi\gamma\rho' h^2 (\sec\alpha - 1)$ and the component attraction along its axis is $\pi\gamma\rho' h (1 - \cos\alpha)$. For a sector of the cone of angle β with its sharp edge along the axis, these are merely affected by a factor $\beta/2\pi$. The extra gravity at the summit of such a sharp promontory is due mainly to its direct attraction and so differs from that for a flat plate by the factor $\beta/2\pi \,.\, (1 - \cos\alpha)$. Due to this uncompensated plate of ocean by itself of depth h and defect of density ρ' $(= \frac{3}{2})$ the direct defect of gravity is (as *supra*) $\frac{3}{2}\frac{\rho'}{\rho}\frac{h}{a}g$, which is 0·1 in 981 per mile of depth. A conical peak of density ρ' and semi-angle 30° thus would produce an excess half this, diminished however by the compensation below, which is of the order of the fraction $\frac{1 + \cos\alpha}{3}\frac{h}{H}$ of it, so negligible if H, the depth of compensation, is 100 miles. In fact it is only the compensation for sideway excess mass, as it attracts more vertically, that is of sensible direct influence on g. Its sideway attraction deflects the plumb line, and so contributes to the oceanic rise of level, but in a way which must be already included in the potential effect, as *supra*. In a sector of a cylinder also, or for a sloping ocean floor, results may readily be calculated in general illustration. It is, of course, the transverse component of local attraction that affects levelling, while pendulums are affected by its vertical component. Theoretically the two are interconnected through the potential, as above indicated in general terms, in a manner illustrated by precise formulae for the case of a set of parallel mountain chains first (1867) in Thomson and Tait's *Natural Philosophy*, or more readily in simpler cases.

Generally, the conclusion is that the depression of sea-level at island stations is not adequate by itself to counteract the direct diminution in gravity due to the low density of the attracting water, much less to produce the observed excess. But if the oceanic defect of density is deeply compensated underneath, down to a uniform standard depth of ocean, any local anomaly of gravity should be due simply to the excess mass over that of water up to an ideal level surface of the local land standing in this ocean.

An illustration involving the potential is more recently reported*. It appears that astronomical observation and direct

* *Proc. Washington Acad.* (Jan. 16).

Digitized by Google

measurement show a difference of as much as 1 mile in the distance between two stations on the north and south coasts of Porto Rico about 35 miles apart. Thus over the island the radius of a level surface is less than the Earth's radius R by 1 part in 35. Thus the expression for the potential locally is

$$V = -g\,(z - x^2/2.\tfrac{34}{35}R + pxz + qz^2 + \ldots),$$

where p and q are to be determined from observed local gradient of g, as to which levelling alone does not give information. Discussions of this kind go back to Bouguer for the Andes, and Cavendish and Maskelyne for Schiehallion*.

* Cf. Cavendish, *Scientific Papers*, vol. II, pp. 402–7.

Digitized by Google

A Contribution to the Theory of Ferromagnetism. By G. S. MAHAJANI, B.A., St John's College.

[*Received* 13 February 1926.]

The following gives a brief summary of a paper, which it is hoped to publish *in extenso* later. It is a contribution to the theory of ferromagnetic crystals, and contains in particular a theoretical explanation of Webster's* experimental results.

Introduction. The fundamental fact of interest with regard to experiments on magnetic substances is that the direction (ψ) of magnetization (I) does not in general coincide with the direction (ϕ) of the external, applied field (H). Experiments on iron crystals have now revealed the fact that this "deviation-effect" ($\phi - \psi$) exists even in crystals of cubic structure and a symmetrical boundary. In other words, cubical symmetry of structure does not make a body magnetically isotropic, as was erroneously assumed by Faraday, Tyndall and Plücker. To account for this "deviation-effect", and in general to relate the two directions (ϕ, ψ) with each other, must certainly be the first object of any theory bearing on this subject. Accordingly Weiss has developed a "macroscopic" theory by introducing the conception of the "molecular field". The idea served many useful purposes, especially in connection with experiments on Pyrhotite. But, as is shown in the present paper, it is quite inadequate to explain the phenomena in the case of crystals with cubic structure.

It seemed, therefore, better to give up the idea of the "molecular field", and start right from the beginning with atomic actions. This alternative was indeed conceived but not adopted by both Weber and Maxwell. It is quite possible that, apart from the external field (H), the internal field which governs the magnetization of the body arises partly from the mutual actions of the "molecular magnets" and partly from some other force, of which we do not know the origin. The rational way of proceeding, however, is to try to see whether we cannot do with only the mutual actions unaided by a force of an obscure origin. This is the point of view adopted in this investigation.

General method. Two types of ferromagnetic crystals are considered:

1. Iron, which has a cubic structure, space-centred cubic, i.e. consisting of two simple cubic lattices, interlocked in such a way that the lattice-points of either occupy the centres of cells of the other.

* "Magnetic Properties of Iron Crystals", by W. L. Webster, *Proc. Roy. Soc.* A, 107 (1925).

Digitized by Google

2. Pyrhotite, which has a non-cubic structure. Two orthorhombic simple lattices are interlocked in a "space-centred" manner, and give on the whole a hexagonal symmetry.

The ultimate particles that form the lattice-points of crystals are supposed to derive their magnetic properties from the presence of electronic orbits. A parameter (a) is introduced to denote the radius of an orbit. The magnetic energy due to mutual actions of the particles is calculated as a series of terms involving third, fifth, seventh, ..., etc. powers of ($1/s$), where s is the lattice-constant. The parameter (a), the radius of an electronic orbit, is also naturally involved. The total magnetic energy of the system in an external field (H) is determined per unit volume; and it is made stationary to search for the steady states of magnetization.

In either of the two types considered the structure consists of two simple lattices. In calculating the internal magnetic energy, it is assumed that the "magnetic elements" (or, the axes of the electronic orbits in our case) arrange themselves, in any steady state, into two groups with two different directions. The two groups severally belong to the two component simple lattices. In saturation states, the two directions coincide; while in the non-magnetized state of the substance they are opposite to each other.

The two distinctive features of the theory in this paper are the association of two distinct directions with magnetic elements in the two component simple lattices, and the replacement of the usual "doublet-conception" by the one of electronic orbits.

Notation:

O-xyz	Coincide with three edges of a cell of either lattice.
(l, m, n)	Direction of I, magnetization.
(l', m', n')	Direction of H, applied field.
ω^3	Volume of a cell.
L_1, L_2	The two component lattices.
$1/\omega^3$	Density of magnetic elements in either lattice; and
$2/\omega^3$	Density in the crystal as a whole.

The work is divided into three parts, and may now be summarized as follows:

Part I.

This deals with cubic structures with saturation states and is devoted mainly to explaining Webster's results. A space-centred cubic structure is specified by

$$\left.\begin{aligned} x &= ps \\ y &= qs \\ z &= rs \end{aligned}\right\} p = q = r \pmod 2,$$

where $2s$ is the side of an elementary cubic cell.

Digitized by Google

The mutual potential energy of two magnetic elements, if identified with two parallel electronic equal orbits, radius a, is $2w$, where

$$w = -\frac{\mu^2}{r^3} P_2 + 3\frac{\mu^2 a^2}{r^5} P_4 *$$

(neglecting higher terms) and

r = distance between the centres,

$\mu = \pi a^2 i$,

$P_2, P_4 = P_2(\cos\theta), P_4(\cos\theta), \ldots,$

and θ = angle between r and the axis of an orbit.

The contribution to the internal energy due to any one magnetic element, say the one at the origin, in the presence of all others at the other lattice-points is

$$W_2 = \sum'_{p,q,r} w_2 = -\sum'_{p,q,r}\left(\frac{\mu^2}{r^3} P_2 - \frac{3\mu^2 a^2}{r^5} P_4\right),$$
$$p = q = r \pmod 2,$$

and the accent indicates that $p = q = r = 0$ is a combination to be omitted.

It is easily seen that

$$r^2 = (p^2 + q^2 + r^2)\, s^2, \qquad \cos\theta = \frac{lp + mq + nr}{(p^2 + q^2 + r^2)^{\frac{1}{2}}}.$$

It is shown in Part III that for cubic structures the third order terms which lead to divergent series involve no variable part of the energy and we are left with

$$W_2' = \text{effective part of } W_2 = +\sum'_{p,q,r} \frac{3\mu^2 a^2}{r^5} P_4(\cos\theta)$$
$$= +\sum'_{p,q,r} \frac{3\mu^2 a^2}{(p^2 + q^2 + r^2)^{\frac{5}{2}} . s^5} P_4(\cos\theta), \quad p = q = r \pmod 2.$$

This reduces, omitting the non-variable part (i.e. not depending upon lmn), to

$$\frac{105}{4}\frac{\mu^2 a^2}{s^5}(3Q - P)(l^2m^2 + m^2n^2 + n^2l^2),$$

where

$$P = \sum'_{p,q,r} \frac{p^4}{(p^2 + q^2 + r^2)^{\frac{9}{2}}}$$
$$Q = \sum_{p,q,r} \frac{p^2q^2}{(p^2 + q^2 + r^2)^{\frac{9}{2}}}$$
$$\Big\} \; p = q = r \pmod 2.$$

* The mutual potential energy of two linear magnets is, however,

$$-\frac{2\mu^2}{r^3} P_2 - \frac{2\mu^2 a^2}{r^5} P_4, \text{ etc.}$$

The second term differs in sign. Actually experimental results require (+) sign. Hence the orbit-conception.

Digitized by Google

Actual calculation shows that

$$0{\cdot}15 < P < 0{\cdot}18,$$
$$0{\cdot}068 < Q < 0{\cdot}076,$$
$$0{\cdot}032 < 3Q - P < 0{\cdot}064.$$

Next, the energy due to an external field (H) is

$$-\mu H\,(ll' + mm' + nn').$$

Therefore the total energy density per unit volume is

$$E = \frac{2\mu^2}{\omega^3}\left[\alpha\,(l^2m^2 + m^2n^2 + n^2l^2) - \frac{H}{\mu}\,(ll' + mm' + nn')\right],$$

where

$$\alpha = \frac{105}{4}\,\frac{a^2}{s^5}\,(3Q - P).$$

(This formula does not take account of temperature effects.)

Let now

$$\lambda x + \mu y + \nu z = 0$$

be a symmetry plane of the crystal structure, and consider a circular plate of the crystal cut parallel to this plane. If H be applied in the plane of the disk, we have

$$\lambda l' + \mu m' + \nu n' = 0.$$

By symmetry of the structure the vector of magnetization will also lie in the plane of the disk and therefore

$$\lambda l + \mu m + \nu n = 0.$$

Making now E stationary subject to

$$\Sigma l^2 = 1, \qquad \Sigma\lambda l = 0,$$

we get

$$\begin{vmatrix} l^3 & m^3 & n^3 \\ l & m & n \\ \lambda & \mu & \nu \end{vmatrix} \div \begin{vmatrix} l' & m' & n' \\ l & m & n \\ \lambda & \mu & \nu \end{vmatrix} = -\frac{H}{2\mu\alpha}.$$

This is a quite general result and includes all the cases considered.

Obviously (H) and (I) will coincide in direction, i.e. the deviation-effect will vanish, when $(l', m', n') = (l, m, n)$, or

$$\begin{vmatrix} l^3 & m^3 & n^3 \\ l & m & n \\ \lambda & \mu & \nu \end{vmatrix} = 0.$$

This result coupled with

$$\Sigma l^2 = 1, \qquad \Sigma\lambda l = 0,$$

gives the directions in the symmetry plane $\lambda x + \mu y + \nu z = 0$, along which the deviation-effect vanishes.

Digitized by Google

Particular cases.

(1) *Circular plate cut parallel to the face of an octahedron:*

$$\lambda = \mu = \nu = 1/\sqrt{3}.$$

In this case $\Sigma l = 0$ makes

$$\begin{vmatrix} l^3 & m^3 & n^3 \\ l & m & n \\ 1 & 1 & 1 \end{vmatrix} = 0$$

an identity, and thus the following experimental fact is explained:

"The amplitude of the waves of σ_n, i.e. the component of magnetization perpendicular to the direction of H, vanishes in stronger fields almost completely*."

(2) *Circular plate cut parallel to the face of a rhombic dodecahedron.*

Here $\lambda = 1/\sqrt{2}, \quad \mu = 1/\sqrt{2}, \quad \nu = 0.$

Solving

$$\left.\begin{vmatrix} l^3 & m^3 & n^3 \\ l & m & n \\ 1 & 1 & 0 \end{vmatrix} = 0 \right\}, \qquad \left. \Sigma l^2 = 1, \quad l + m = 0 \right\}$$

we get exactly the same result as required by experimental data:

"...The normal components σ_n show four zeros within 180° in the direction of the axes at 0, 55°, 90°, 125°, and 180°†."

Strictly speaking, theory yields not 55° and 125°, but $\cos^{-1}(1/\sqrt{3})$ and $\pi - \cos^{-1}(1/\sqrt{3})$, that is 54° 44′ and 125° 16′.

(3) *Circular plate cut parallel to the face of the cube.* This is the case of Webster's experiments. We have $\lambda = \mu = 0$, $\nu = 1$, and therefore $n = n' = 0$. The result reduces to

$$\frac{\begin{vmatrix} l^3 & m^3 \\ l & m \end{vmatrix}}{\begin{vmatrix} l' & m' \\ l & m \end{vmatrix}} = -\frac{H}{2\mu\alpha}.$$

Putting

$$\left.\begin{aligned} (l, m) &= (\cos\psi, \sin\psi) \\ (l', m') &= (\cos\phi, \sin\phi) \end{aligned}\right\},$$

we get

$$\frac{\sin(\phi - \psi)}{\sin\psi\cos\psi\cos 2\psi} = \frac{2\alpha\mu}{H}.$$

* Kunz, *Bull. Nat. Res. Council*, vol. 3, Part 3, p. 180 (1922).
† *Ibid.*

Digitized by Google

Further, if we write $\mathfrak{I}$ instead of I, to denote that we are dealing with saturation states, we see that

$$\frac{\mathfrak{I}_n}{\sin\psi\cos\psi\cos 2\psi} = \frac{2a\mu\mathfrak{I}}{H},$$

i.e. $$(\mathfrak{I}_n)_{\max.} H = \tfrac{1}{2}a\mu\mathfrak{I} = \text{constant} \qquad \ldots\ldots(1).$$

The explanation of all the experimental curves of Webster is contained in the two formulae

$$\left.\begin{aligned} \frac{H\sin(\phi-\psi)}{\mathfrak{I}\sin\psi\cos\psi\cos 2\psi} = \frac{2a\mu}{\mathfrak{I}} = a.\omega^3 \\ (\mathfrak{I}_n)_{\max.} H = \tfrac{1}{2}a\mu\mathfrak{I} \end{aligned}\right\}.$$

Quantitatively, the second of the results above is made to fit Webster's Graph 4, by adjusting a. All the quantities in equation (1) are known, or can be calculated except the parameter a, the radius of an electronic orbit. Thus the proper adjustment of a to fit the experimental curves yields the magnitude of a.

The results are:

$$\left.\begin{aligned} &\text{Disk } A \quad 0{\cdot}212 > a \times 10^8 > 0{\cdot}147 \\ &\text{Disk } B \quad 0{\cdot}331 > a \times 10^8 > 0{\cdot}232 \end{aligned}\right\}.$$

The conclusion is reached that

$$a = O\,[10^{-9}].$$

The molecular field. Our work shows that if we start from the bulk-conception of the molecular field, we must adopt a "third-power" law and not a "fourth-power" one, as is empirically suggested by Webster, in the case of cubic structures. Incidentally a better generalization of Weiss's law of simple proportionality in the case of non-cubic structures is suggested, which includes both the cubic and the non-cubic cases. Thus in place of Weiss's

$$(H_m)_\mu = N_\mu I_\mu,$$

we write $$(H_m)_\mu = N_\mu\,[a_1 I_\mu + a_3 I_\mu{}^3 + a_5 I_\mu{}^5 + \ldots].$$

Adopting this form, we obtain in the case of a circular plate the following result:

$$\frac{H\begin{vmatrix} l & m \\ l' & m' \end{vmatrix}}{Ilm\,[a_1(N_x - N_y) + a_3(l^2N_x - m^2N_y)\,I^2 + a_5(l^4N_x - m^4N_y)\,I^4 + \ldots]} = 1.$$

In the non-cubic case $N_x \neq N_y$.

Digitized by Google

Retaining only the first non-vanishing term in the denominator gives

$$\frac{H\begin{vmatrix} l & m \\ l' & m' \end{vmatrix}}{Ilm} = a_1 (N_x - N_y).$$

Take $a_1 = 1$ and we obtain exactly Weiss's result

$$\frac{H \sin(\phi - \psi)}{I \sin\psi \cos\psi} = (N_x - N_y).$$

On the other hand, in the cubic case

$$N_x = N_y = N, \text{ say.}$$

Therefore retaining only the first non-vanishing term gives

$$\frac{H\begin{vmatrix} l & m \\ l' & m' \end{vmatrix}}{I^3 lm\,(l^2 - m^2)} = a_3 N,$$

i.e. in the saturation case

$$\frac{H \sin(\phi - \psi)}{\mathfrak{I} \sin\psi \cos\psi \cos 2\psi} = a_3 N \mathfrak{I}^2,$$

which agrees with the result we obtained by the consideration of "atomic actions".

Quantitatively, the agreement is fairly satisfactory. We obtain the following result:

	Present theory	Webster's experiments
M_B	576	620
M_A	368	470

where M_B, M_A denote the magnitude in gauss of the molecular field in the case of saturation when H acts along one of the structural axes, for disks B and A respectively.

Part II.

This deals with non-cubic structures. Two methods are pointed out for dealing with the "third-order" terms in the energy function. It is shown that, in the cubic structures, these involve no variable part of the energy; while in the non-cubic cases, these involve the most effective part of the variable energy. Hence the simple doublet conception suffices in this case. The same results are arrived at as Weiss's.

Digitized by Google

Part III.

In this, an attempt is made to extend the theory to non-saturation states. The two general results that are obtained may be stated here. Cubic case:

$$\frac{H\cos(\phi-\psi)}{I\left[2q-\frac{a^2}{8}\{4(h-3j)\cos 4\psi.f_1+2h'(1+\cos^2 2\psi)+3j'(1-\cos 4\psi)\}\right]}$$
$$=-\tfrac{1}{2}\omega^3,$$

and
$$\frac{I-H\sin(\phi-\psi)}{\sin 4\psi\,[(hf_2+h'f_1)-3(jf_2+j'f_1)]}=-\frac{a^2}{16}\mathfrak{I}^2\omega^3,$$

where q, h, h', j, j' are constants defined by the structure thus:

$$\left.\begin{aligned} h&=\sum_{p\ q,r}\frac{\partial^4}{\partial x^4}\left(\frac{1}{r}\right)\\ j&=\sum_{p,q,r}\frac{\partial^4}{\partial x^2\partial y^2}\left(\frac{1}{r}\right)\end{aligned}\right\}\ (p,q,r)\ \textit{all even},$$

$$\left.\begin{aligned} h'&=\sum_{p,q,r}\frac{\partial^4}{\partial x^4}\left(\frac{1}{r}\right)\\ j'&=\sum_{p,q,r}\frac{\partial^4}{\partial x^2\partial y^2}\left(\frac{1}{r}\right)\end{aligned}\right\}\ (p,q,r)\ \textit{all odd},$$

$$r^2=(p^2+q^2+r^2)\,s^2,$$

$$f_n=f_n(I/\mathfrak{I})=\cos n(\psi_1-\psi_2),$$

ψ_1 and ψ_2 being the directions of the magnetic elements in the two component simple lattices. The constant q is rather too complicated to define in this summary.

Formulae are obtained for the non-cubic structures also, in non-saturated states.

Concluding remarks. The theory requires to be completed in three directions:

1. To consider which of the steady states are stable. This will probably bear on the hysteresis phenomena.

2. To consider the effect of temperature.

3. To put the formulae for the third order terms in Part II in a form suitable for numerical computation, and thus to test the theory in the case of non-cubic structures quantitatively.

The writer wishes to thank Prof. C. G. Darwin, Mr A. E. Ingham, Dr Kapitza, Mr Webster and *especially* Mr E. Cunningham for their interest in the work.

Digitized by Google

Over-tension in a Condenser Battery during a sudden discharge.
By P. KAPITZA, Ph.D. Fellow of Trinity College, Assistant Director of Magnetic Research at the Cavendish Laboratory, Clerk Maxwell Student of Cambridge University.

[*Read* 18 January 1926.]

(Plate II.)

In the recent experiments on the Zeeman effect in strong magnetic fields* made by Mr H. W. B. Skinner and myself for obtaining an intense source of light for a small fraction of a second, we have been discharging a condenser battery consisting of 32 Leyden jars connected in parallel with a capacity of one-tenth of a microfarad. The discharge was made through a small spark gap by means of an oil immersed switch (the details of this arrangement are to be found in the above-mentioned paper). When the discharge was produced a curious phenomenon was sometimes observed. After the battery had been charged to its maximum tension, and the switch had been put into operation, only a small spark occurred in the spark gap, and the main part of the discharge went over the top of the insulators of one of the Leyden jars.

By taking the photograph of the spark on a film moving rapidly before a slit which was illuminated by the spark, it was possible to see the details of the phenomenon.

In Figs. 1, 2 and 3 are shown three photographs taken with different self-inductions in series with the spark gap. On Photograph 1 a normal discharge is seen in which no flash over the insulation occurs in the condenser battery. About 20 full oscillations can be seen, after which the spark is extinguished. Photographs 2 and 3 represent the spark in the case when the discharge flashes over one of the condensers. From these photographs it can be seen that in the case of Photograph 2 the spark is extinguished after the first half of the oscillation, as only one line is seen, and that in the case of Photograph 3 it is extinguished after the full oscillation, as two lines are seen. This shows the very peculiar character of the over-tension occurring during a sudden discharge in a condenser battery. In this paper a short account of these phenomena will be given in which it will be shown how it is possible to account for these over-tensions, and how they can be avoided.

From general considerations it can be seen that during the discharge of a condenser battery, when the self-induction of the

* *Proc. Roy. Soc.* A, vol. 109, p. 225 (1925).

Digitized by Google

Fig. 1

Fig. 2

Fig. 3

Digitized by Google

Digitized by Google

connecting leads together with the self-induction of the Leyden jars is of the same order as the self-induction of the spark circuit, electrical oscillations can be set up between the Leyden jars which, being superimposed on the tension of the general discharge, may produce over-tension in separate Leyden jars.

Let us consider a large number of Leyden jars connected in series as shown on Fig. 4.

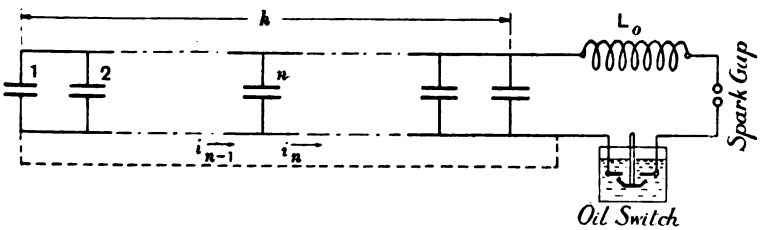


Fig. 4

Let the capacity of each of the Leyden jars be c; the self-induction of the connecting leads l; the current in the connecting leads of the Leyden jars i_n, i_{n-1}, and the tension of the Leyden jar e_n. We then get a series of equations of the following type:

$$i_n - i_{n-1} = -c\frac{de_n}{dt} \qquad \text{.....(1),}$$

$$e_n - e_{n-1} = -l\frac{di_{n-1}}{dt} \qquad \text{.....(2).}$$

The number of these equations is twice the number of the jars, and it is rather difficult to treat the system of these differential equations for a large number of Leyden jars, but with a quite sufficient approximation we can replace them with two partial differential equations. Taking the capacity and the self-induction as equally distributed along the condenser battery we get

$$\left.\begin{aligned} c &= C\,dx \\ l &= L\,dx \end{aligned}\right\} \qquad \text{.....(3),}$$

$$i_n - i_{n-1} = di,$$

$$e_n - e_{n-1} = de.$$

Taking the axis of coordinates along the battery as shown in Fig. 4, we obtain the equations

$$\frac{di}{dx} = -C\frac{de}{dt} \qquad \text{.....(4),}$$

$$\frac{de}{dx} = -L\frac{di}{dt} \qquad \text{.....(5).}$$

Digitized by Google

Now to obtain a solution of these equations it is necessary to satisfy the following initial conditions: that when

$$\left.\begin{array}{l} t = 0, \\ \text{(A)}\ \ i = 0 \\ \text{(B)}\ \ e = e_0 \end{array}\right\} \quad (0 \leqq x \leqq h),$$

where e_0 is the potential to which the battery is charged and h may be called the total length of our condenser battery. We must also satisfy the following conditions: that

$$\left.\begin{array}{lll} \text{(C)} & \text{at } x = 0, & i = 0 \\ \text{(D)} & \text{at } x = h, & e = L_0 \dfrac{di}{dt} \end{array}\right\} \quad (any)\ t,$$

where L_0 is the outside self-induction.

The most general solution for i and e in these equations can be shown to be of the following form:

$$i, e = \left.\begin{array}{l} \Sigma\,(A \sin mx + B \cos mx) \sin amt + \\ + \Sigma\,(A' \sin mx + B' \cos mx) \cos amt \end{array}\right\} \quad \ldots\ldots(6),$$

where

$$a = \sqrt{\frac{1}{CL}} \quad \ldots\ldots(7).$$

It is easily seen that in satisfying equations (4) and (5) and conditions (A) and (C), we obtain the following expressions for i and e:

$$i = \Sigma A \sin mx \sin amt \quad \ldots\ldots(8),$$

$$e = La\Sigma A \cos mx \cos amt \quad \ldots\ldots(9).$$

The boundary condition (D) requires that

$$L \cos mh = L_0 m \sin mh \quad \ldots\ldots(10).$$

This gives the value for m. If, for brevity, we put

$$\mu_m = mh \quad \ldots\ldots(11),$$

and

$$\gamma = \frac{hL}{L_0} \quad \ldots\ldots(12),$$

we obtain from (10) the following transcendental equation

$$\gamma = \mu_m \tan \mu_m \quad \ldots\ldots(13),$$

the roots of which determine the value of m. Introducing these roots in the general expressions for e and i we obtain

$$i = \Sigma A_m \sin \frac{\mu_m x}{h} \sin \frac{\mu_m at}{h} \quad \ldots\ldots(14),$$

$$e = aL\Sigma A_m \cos \frac{\mu_m x}{h} \cos \frac{\mu_m at}{h} \quad \ldots\ldots(15).$$

Digitized by Google

Now the coefficient A_m will be determined from the initial condition (B). It has been shown by Poisson* that expansions in Fourier series with coefficients which are roots of transcendental equations such as (13) can be treated similarly to ordinary Fourier series.

We find $$A_m = \frac{e_0}{La}\frac{1}{N_m}\int_0^h \cos\frac{\mu_m x}{h}\,dx \qquad \text{.....(16),}$$

where $$N_m = \int_0^h \cos^2\frac{\mu_m x}{h}\,dx \qquad \text{.....(17),}$$

and we obtain $$A_m = \frac{4e_0}{La}\frac{\sin\mu_m}{2\mu_m + \sin 2\mu_m} \qquad \text{.....(18).}$$

The final solution in our case is therefore

$$e = 4e_0 \sum_0^\infty \frac{\sin\mu_m}{2\mu_m + \sin 2\mu_m}\cos\frac{\mu_m}{h}x\cos\mu_m\frac{a}{h}t \qquad \text{.....(19),}$$

$$i = \frac{4e_0}{La}\sum_0^\infty \frac{\sin\mu_m}{2\mu_m + \sin 2\mu_m}\sin\frac{\mu_m}{h}x\sin\mu_m\frac{a}{h}t \qquad \text{.....(20).}$$

Let us first consider the case when the outside self-induction L_0 is very large. It is easily seen that γ in equation (13) represents the ratio of the total internal self-induction hL to the outside self-induction L_0. When γ is very small we find that the roots of the equation are close to $n\pi$, where n is any positive integer from 0 to ∞. In this case all the A's vanish except $A_0 = \frac{1}{4}$, and if we take into consideration that the smallest root of the equation (13) is

$$\mu_0 = \sqrt{\frac{hL}{L_0}} \qquad \text{.....(21),}$$

we easily obtain from (19) and (20), taking into consideration (7),

$$e = e_0\cos\frac{t}{\sqrt{ChL_0}} \qquad \text{.....(22),}$$

$$i = e_0\sqrt{\frac{Ch}{L_0}}\frac{x}{h}\sin\frac{t}{\sqrt{ChL_0}} \qquad \text{.....(23).}$$

If we remember that Ch is the total capacity of the battery, we see that with a large outside self-induction there are, in the normal way, no oscillations in the battery during the discharge, as e is independent of x, and the discharge is of the same type as would

* *Journal de la Polytechnique*, cahier 18: Poisson, "Mémoire sur la manière d'exprimer les fonctions en série de quantités périodiques". In this work Poisson treats a problem very similar to ours. He calculates the vibrations of a load attached to a bar with an equally distributed mass. The problem dealt with in my case is very like that treated by Poisson, as the self-induction of the outside circuit is equivalent to the inertia of the load, and the distributed capacity and self-induction of the condensers correspond to the distributed mass and to the elasticity of the load.

Digitized by Google

be obtained with a condenser of a capacity Ch, with an outside self-induction L. From (23) we see that the current i increases linearly along the battery.

If γ be large, the tension along the condenser will not be equally distributed during the discharge, and from expression (19) we can easily see that the maximum value of e will be at $x = 0$, the end of the condenser opposite to the end from which the current is taken. e in this place will equal

$$e = e_0 \Sigma \frac{4 \sin \mu_m}{2\mu_m + \sin 2\mu_m} \cos \frac{\mu_m at}{h} \qquad \text{.....(24).}$$

From the time $t = 0$ to the time $t = \frac{h}{a}$ the tension at this place is constant, as can at once be seen from the way in which we determine the coefficient A_m, (16) and (17). Further on it decreases to zero and then increases again with opposite sign. In order to see what really happens we have to calculate the value of this sum for a longer time than $t = \frac{a}{h}$ and for different γ's. This is not difficult as it is quite sufficient to calculate four terms of this expression which approximately converges as $\frac{1}{m^2}$. If $\gamma = 1$ we find that after the first half of the oscillation the tension at $x = 0$ will be 1·28 of the initial one. If $\gamma = 2$ we find that after $1\frac{1}{2}$ full oscillations the tension at the same place will be equal to 1·35 of the initial one. If $\gamma = 4$ after ten full oscillations the tension at $x = 0$ will be 1·56 of the initial one. This shows that with increasing γ the over-tension will gradually increase, but the maximum will occur later. This explains the fact that sometimes the breaking down of the condenser happens only after several half-oscillations. Calculating the terms in this series it is easily seen that the first two terms are the most important ones. At the beginning they have opposite signs, but after some time of oscillation they may have the same sign and maximum amplitudes together. This moment will give the over-tension. If the first root of equation (13) is μ_0, and the second μ_1, and the ratio $\frac{\mu_1}{\mu_0} = k$, then the over-tension will happen after n half-oscillations where n is defined by

$$kn\pi - n\pi = m\pi \qquad \text{.....(25),}$$

and m is any odd number. From this formula (25) we find that the ratio of the first and second roots of this equation has to be 4 in order to get the over-tension after the first half-oscillation, $3\frac{1}{2}$ to get it after the second half-oscillation, $3\frac{1}{3}$ to get it after the third oscillation and so on.

In these calculations we have not taken into consideration the

Digitized by Google

damping of the oscillations due to the resistance in the circuit. This fact, indeed, makes it impossible for the over-tension to occur very late during the discharge, but for the first two or three cycles it is probably permissible to ignore the resistance, and this very much simplifies the analysis.

From the analysis it is also seen that in general the over-tension during the discharge of a condenser battery could easily be avoided if the sparking circuit was connected to different ends of the condenser battery, as is shown by a dotted line on Fig. 4. In this case each Leyden jar when discharging has the same self-induction in each circuit, and when all the Leyden jars discharge with the same period, no oscillations between the jars can occur, as there is no difference of potential between them. This also applies to a separate Leyden jar which, in order to avoid electrical over-tension, must be connected in such a way that the leads are taken from opposite ends. It may also be mentioned that, as can be seen from (20), some irregularity must be observed in the current, as it is expressed by an equation similar to (19) for the tension. On Photograph 1 it can be observed that the intensity of the spark varies in an irregular way from one oscillation to another, and this is probably due to variations in the current.

In certain cases in practice it is possible to use oscillations in Leyden jars in order to obtain very short and sharp impulses of the electrical current. For this purpose it is only necessary to place a spark gap at the opposite end of the condenser battery, connecting this spark gap to the condenser battery without any self-induction, and adjusting it so as to make it slightly larger in size than the spark gap in the circuit with the oil switch. By adjusting the self-induction in the series with the main spark gap, it will be possible to produce over-tension in the subsidiary spark gap after one, two or more oscillations, and as the discharge will at this moment pass in the subsidiary spark gap, we shall only obtain one or two oscillations in the main spark gap.

In conclusion I would like to express my thanks to Prof. A. N. Kryloff, with whom I have discussed the mathematical part of this paper.

Digitized by Google

The Period of Decay of Radium B and Radium C. By P. BRACELIN. (Communicated by Dr J. CHADWICK.)

[*Read* 18 January 1926.]

In determining the transformation constant of a radio-active substance by direct experiment, two general methods are available, depending on the precision possible in individual measurements. In the first case, a very accurate determination of the activity of the substance is made at two times, the length of time that elapses between the measurements being relatively unimportant as long as it can be estimated exactly. Alternatively, the decay of the activity may be followed by measurements not so precise individually, and in this case the accuracy of the result is largely influenced by the length of time over which the observations are extended. This requires, over even a small number of half-value periods, that the method of measurement employed is capable of the accurate comparison of widely different activities.

It is a defect in the electroscope method of comparing different ionisation currents that the accuracy obtainable falls off greatly as the ratio of the currents increases, and the error introduced is difficult to allow for. This objection does not apply to galvanometer measurements where, by the use of suitable shunts, it is possible to obtain an accurate comparison of very different intensities of current. The actual strength of the currents measurable even by a sensitive galvanometer demands of course the use of much greater quantities of radio-active material, but, where this is possible, the method seems to be less open to criticism than the use of the electroscope.

In the experiments described below the galvanometer method has been applied to the case of radium C and has been found to give very consistent results. Similar measurements of the decay of radium active deposit have enabled an estimate of the period of radium B to be made.

The apparatus consisted essentially of a parallel plate condenser, one plate being connected through a galvanometer and a megohm resistance to earth, and the other to a source of potential. The plates were enclosed in a glass vessel in communication with a drying tube, a manometer, and an oil pump. Thus dry air at any desired pressure could be obtained, and standard conditions for the different experiments ensured.

The radio-active source was fixed on one of the plates, and the saturation current measured by the galvanometer using a telescope and scale arrangement, which allowed deflections to be estimated

Digitized by Google

within ·2 mm. with accuracy. The scale was carefully calibrated by means of known currents during each experiment. The resistances of the galvanometer and the shunts employed in conjunction with it were determined by a Post Office box guaranteed correct to 1 in 1000.

The attainment of saturation of α-particle ionisation currents is a difficult matter at atmospheric pressure, but is rendered much easier by working at reduced pressures. Accordingly the experiments were carried out at pressures from 20 to 40 cm. of mercury. In order to be certain that the measurements were performed within the saturation range, the applied potential was varied over the first few readings, starting at a considerably lower value than was afterwards maintained. As these are also the largest currents, this ensures that all the readings are taken at saturation, the applied fields being well below the limit where ionisation by collision was possible. Preliminary experiments with polonium showed also that over a wide range of applied potentials, when saturation was attained, the galvanometer deflection was very free from fluctuations, indicating that the single reading of a particular activity, which it is possible to take when following a decaying ionisation current, was an accurate index of the actual saturation value of the current at that instant.

Radium C. The radium C sources were prepared in the ordinary way by spinning a nickel disc in a solution of radium active deposit in hydrochloric acid at 80° C. The active deposit after removal from the emanation was left for 40 minutes before solution. By this time the radium A present, which would also be deposited by this process on the nickel disc, had decayed to about ·01 per cent. of the original amount. The quantities of radium C used showed an initial activity of the order of 20 mgms. radium, and with these it was possible to follow the decay for a little less than two hours, i.e. six half-value periods.

It is important to check the purity of the sources used, and this was accomplished by measuring the decay over an hour with an α-ray electroscope, five hours after the radium C was collected from solution. It was estimated that 0·5 per cent. radium A deposited initially would cause an increase of 0·2 minute in the observed period over the first two hours, while from 5–6 hours afterwards the difference from the real period would amount to 1·5 minutes. This would easily be detectable by the electroscope measurements. In fact, one experiment in which the RaC was collected about 20 minutes after removal of the deposit from the emanation enabled this to be verified, the period from the galvanometer measurements being estimated at 19·9 minutes and from the electroscope readings as 21·4 minutes. In the experiments from which the transformation

Digitized by Google

constant was obtained, the electroscope measurements showed that no perceptible change in the decay period had taken place, giving good grounds for the belief that pure radium C was being used.

The most probable values of the transformation constant λ were calculated by the method of least squares from the results of four experiments, and are given below with their probable errors:

$\lambda \times 10^2$ min.$^{-1}$
$3{\cdot}516 \pm {\cdot}002$
$3{\cdot}521 \pm {\cdot}003$
$3{\cdot}507 \pm {\cdot}003$
$3{\cdot}511 \pm {\cdot}002$

giving a mean value for $\lambda = (3{\cdot}514 \pm {\cdot}007) \times 10^{-2}$ min.$^{-1}$ The corresponding period is $T = 19{\cdot}72 \pm {\cdot}04$ minutes. This is about 1 per cent. greater than the generally accepted value of 19·5 minutes as determined by F. v. Lerch*.

Radium B. The sources of radium active deposit were obtained by exposing a brass disc to decaying radium emanation for about 15 hours, without the application of an electric field. The latter, although of value in concentrating the deposit on the collecting disc, has been found to affect in a variable manner the relative amounts of RaA, RaB, and RaC deposited, and thus to render inaccurate the calculation of the resulting decay based on the assumption of radio-active equilibrium. Three experiments were performed in each of which the decay was followed over nearly three hours.

Under the above conditions, transient equilibrium is attained after about five hours' exposure to the emanation, and after removal of the latter, the amount of radium C present at any instant is given by the formula†

$$\frac{\text{Amount of RaC at time } t}{\text{Initial amount of RaC}} = -(\lambda_2 - \lambda_1)(\lambda_3 - \lambda_1)(\lambda_4 - \lambda_1) \sum_2^4 \frac{e^{-\lambda_s t}}{(\lambda_1 - \lambda_2)(\lambda_3 - \lambda_2)(\lambda_4 - \lambda_2)},$$

where λ_1, λ_2, λ_3, λ_4 are the transformation constants of radium emanation A, B, C respectively.

The initial decay is complicated by the presence of radium A, which also gives α-particles, but after about 30 minutes the effect of the radium A is negligible, and the observed activity ought to be in agreement with the values calculated from the above equation.

The accepted values for λ_1 and λ_2 were taken and the value $3{\cdot}514 \times 10^{-2}$ min.$^{-1}$ for λ_4 as found above. For λ_3 various numerical

* F. v. Lerch, *Ann. der Phys.* vol. 20, 2, pp. 345–54 (June, 1906).

† Cf. Rutherford, *Radioactive Substances and their Radiations*, p. 498.

Digitized by Google

values were substituted, and it was found that while other values could be discarded, the values $2{\cdot}587 \times 10^{-2}$ min.$^{-1}$ and

$$2{\cdot}596 \times 10^{-2} \text{ min.}^{-1}$$

for λ_3 both gave very good agreement with the experimental results. If anything, the experimental data seem to indicate that values calculated on the basis of $\lambda_3 = 2{\cdot}596 \times 10^{-2}$ min.$^{-1}$ fall a little too quickly and on $\lambda_3 = 2{\cdot}587 \times 10^{-2}$ min.$^{-1}$, rather too slowly, but the calculated values are in too close agreement over the range of time of the experiments to allow of discrimination between them. These values correspond to half-value periods of 26·7 and 26·8 minutes respectively for radium B, and both have support from previous determinations. F. v. Lerch* using pure RaB found a period of 26·7 minutes, and Mme Curie†, by measuring the decay of the active deposit from 7–10 hours after removal from the emanation, fixed on a period of 26·8 minutes. In conjunction with the above measurements, these determinations suggest that the correct value of the period is intermediate between these two values.

The use of a longer period for radium C in calculating the decay of the active deposit requires that the latter decay more slowly than has formerly been thought. Confirmatory evidence on this point has been found by other workers in this laboratory when checking up deposit sources. This was attributed to the disturbing influence of the electric field used in collecting the sources, but it seems probable that part, at least, of this discrepancy has been

Amount of RaC present.

Time in minutes	Lawson and Hess	From present work
0	100	100
60	50·05	50·23
70	41·21	41·50
80	33·73	33·92
85	30·23	30·60
90	27·10	27·51
100	21·81	22·19
110	17·43	17·79
120	13·96	14·20
130	11·06	11·30
140	8·746	8·955
150	6·868	7·075
160	5·400	5·569
170	4·256	4·379
180	3·342	3·437
190	2·628	2·690
200	2·070	2·107

* F. v. Lerch, *loc. cit.*

† Mme Curie, *Radioactivité*, vol. 2, p. 322 (1910).

Digitized by Google

due to the shorter estimate of the period of radium C in the calculations.

As these sources are frequently employed in radio-active work, it may be of use to give a comparison of values calculated on the above basis with those given in a table by Lawson and Hess* based on previous calculations.

The difference is negligible over the first hour. As Lawson and Hess give values at minute intervals from zero up to 200 minutes, intermediate values may easily be interpolated from the above.

Summary. Measurements of the saturation currents produced between two parallel plates by a source of radium C of carefully tested purity have been performed by an accurate galvanometer method. From the results a transformation constant

$$\lambda = (3{\cdot}514 \pm {\cdot}007) \times 10^{-2} \text{ min.}^{-1}$$

has been assigned to radium C corresponding to a half-value period of $19{\cdot}72 \pm {\cdot}04$ minutes.

Similar measurements with radium active deposit have enabled a value to be estimated for the period of radium B between 26·7 and 26·8 minutes, which is the limit of accuracy of the experimental data.

I wish to express my thanks to Dr J. Chadwick for suggesting this work, and both to him and to Sir Ernest Rutherford for their interest and advice during its prosecution; also to Mr G. R. Crowe for the preparation of the sources used.

* Lawson and Hess, *Wien. Ber.* 2 A, 127, pp. 626–7 (1918).

Digitized by Google

On the Diurnal Variation of Ultra-Short Wave Wireless Transmission.* By Mr E. V. APPLETON, St John's College.

[*Received* 10 March, *read* 3 May 1926.]

During the last few years the intensive study of the transmission of very short waves has revealed to us some remarkable facts relating to their use in long-distance communication. The experimental data have been obtained mainly by wireless amateurs and the Marconi Company in England, and by Major Mauborgne and Dr A. H. Taylor† in America. The low attenuation of such ultra-short waves in long-distance transmission is, of course, well known and need not be emphasized here. But these waves have been found to possess other remarkable properties which make them extremely interesting from a scientific point of view. For example, with wave-lengths of the broadcasting band (300–500 metres) we are familiar with the general, though variable, increase of signal strength at night. But with much shorter waves (e.g. 30 metres) cases often occur in which signals, though heard loudly during the day-time, become inaudible at night. Also, during the day-time it is not usual for a long-wave station to be inaudible at moderate distances from the transmitter and yet easily audible at much greater distances. Yet this phenomenon is quite common when very short wave-lengths are used. In this communication an attempt is made to explain some of the more salient features of such phenomena in terms of the ionized layer theory.

In the first place it must be emphasized that the propagation of very short waves over long distances must be almost wholly due to atmospheric influences, since, for such high frequencies, the diffractive bending round the earth is small and the ground absorption considerable. For example, with a wave-length of 30 metres the effect of the ground ray is very small indeed except near the transmitter and, with increasing distance up to 500 miles, the signal intensity becomes quite negligible. At a distance of about 500 miles‡, however, the signals become suddenly stronger and are detectable up to 1500 miles, beyond which reception becomes uncertain. There is thus a zone of silence round the transmitter with an outer radius of approximately 500 miles. If, in accordance with the theory of

* The term "ultra-short" is here applied to wave-lengths less than the critical band indicated by the magneto-ionic theory (i.e. about 200 metres). It seems desirable, for historical reasons, to retain the term "short waves" for the broadcasting wave-lengths 200–600 metres.

† Taylor, *Proc. Inst. Rad. Eng.* vol. 13 (Dec. 1925).

‡ Day-time transmission is here considered.

Digitized by Google

atmospheric influence which has been shown to explain the phenomena on broadcasting wave-lengths, we interpret the sudden increase of signal intensity at 500 miles as being due to the effect of a ray deviated back to the ground by the upper atmosphere, it has been shown by Mr Barnett and the writer* that an estimate of the minimum number of electrons or ions in the ionized layer can be made. In this connection let us consider the case of refraction by a medium of gradually varying refractive index, the variation being so gradual that the medium may be considered as homogeneous within a few wave-lengths. Let *ABCDEF* be a ray which enters the ionized layer at *B* and emerges from it at *E* (see Fig. 1).

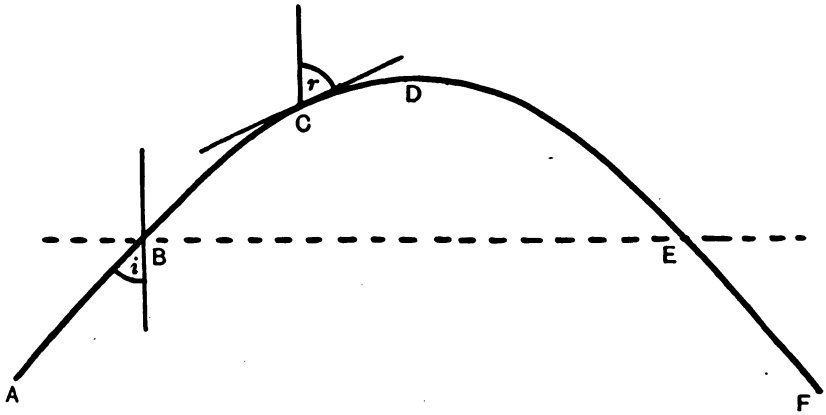


Fig. 1

Let i be the angle of incidence at the lower boundary of the ionized layer and μ_0 the refractive index of unionized air. Then it is easily shown from the general principles of optics that the refractive index μ at any point C is given by the relation

$$\mu_0 \sin i = \mu \sin r \qquad \ldots\ldots(1),$$

where r is the angle between the ray direction at C and the vertical. Now at the highest point D of the trajectory we have $r = \pi/2$ and, since μ_0 may be considered unity,

$$\mu_D = \sin i \qquad \ldots\ldots(2).$$

Now according to the magneto-ionic theory, for wave-lengths which are small compared with the critical band, we may write approximately

$$\mu^2 = 1 - \frac{4\pi N e^2}{m p^2} \qquad \ldots\ldots(3),$$

* Appleton and Barnett, *Proc. Roy. Soc.* A, vol. 109 (1925).

Digitized by Google

where N is the number of electrons (of charge e and mass m) per c.c., and p is the angular frequency of the radiation. From (2) and (3) we deduce that

$$\cos i = 2\sqrt{\frac{\pi N}{m}}\frac{e}{p} \qquad \ldots\ldots(4),$$

so that if i and p are known, N may be estimated. In this way* it has been shown that the value of N is approximately 10^5 per c.c. We may further note that reflection at vertical incidence ($i = 0$) is possible when

$$p = 2e\sqrt{\frac{\pi N}{m}} \qquad \ldots\ldots(5).$$

If the electron concentration is not any greater than 10^5 per c.c. at any height, this would indicate that wave-lengths less than 100 metres would not be reflected at vertical incidence, but would penetrate the layer and leave the earth. If the maximum concentration were 10 times greater, the critical wave-length would be reduced $\sqrt{10}$ times.

Let us now consider the phenomena of ultra-short wave propagation over various distances. Consider a transmitter at T (see Fig. 2) and a receiver R moving successively to R_1, R_2, R_3 and R_4. Experiments show that a receiver near the transmitter, say at R_1, receives audible signals, but, with increasing distance, this signal disappears so that the transmitter is not heard at R_2 and R_3. At R_4, however, the atmospheric ray assumes prominence resulting in signals of appreciable intensity.

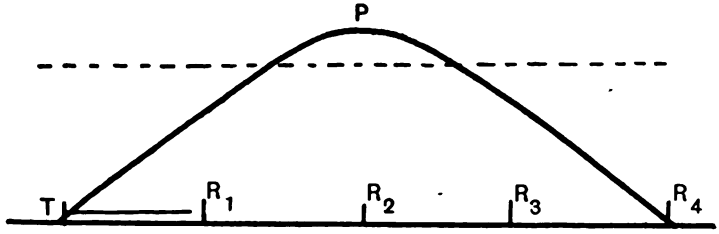


Fig. 2

To explain why the signals at R_4 are stronger than at R_3 we have to consider two possibilities. In the first place it can be shown that the "reflection coefficient" of the ionized layer for such cases as we are considering is greater the greater the angle of incidence, so that it might be argued that the increase of signal strength in

* If the effective carriers were molecular ions, which is, however, unlikely, this number would be increased to 4×10^8 in the case of hydrogen and 3×10^9 in the case of nitrogen.

Digitized by Google

proceeding from R_3 to R_4 is due to the reduced absorption, due to the more nearly grazing incidence. According to this explanation the negligible signals at R_2 and R_3 are due to absorption-limitation. But there is a second possibility. We note from (4) that, as the angle of incidence is increased, fewer electrons are required to bring the ray horizontal and thus back to earth. We therefore can explain the phenomenon if we assume that the ray TPR_4 is a critical one, in that there are just sufficient electrons in the ionized layer to bend it back, while for smaller angles of incidence (and shorter distances) the rays suffer from electron-limitation and escape from the atmosphere altogether. The theory put forward here is that the second explanation is the correct one; that is to say that, in the wireless spectrum we employ wave-lengths which, in certain cases, the atmosphere, due to its finite electronic content, is unable to retain. The reason for the acceptance of the second alternative will now be discussed.

It has been shown that, in the case of the longer wave-lengths of the broadcasting band (300–500 metres), the rays returned from the upper atmosphere are, in general, of greater amplitude at night than during the day. The theory of diurnal variation which has been advanced to explain this is that the ionization, which produces the reduction in refractive index necessary to bring the ray horizontal, exists at lower levels during the day-time than during the night. In this case there is no electron-limitation of bending since the rays are bent down both by day and by night. But, due to the higher level at which bending occurs at night, the absorption resulting from collisional "friction" is smaller than that experienced during the day. Thus, if the first alternative considered in the case of ultra-short waves is correct, and the negligible signal intensity at R_3 is due to the excessive absorption of the ray reaching that point in the day-time, we should expect the absorption to be less at night and a signal strength increase to ensue. The signal strength at R_4 should also increase and the outer radius of the silent zone should be decreased. But in practice no such increases of signal intensity do occur; in fact the signals normally obtained at the edge R_4 of the silent zone disappear, the radius of the zone being increased. The theory of electron-limitation here presented adequately explains this, for if there are only just enough electrons to deviate the ray down in the day-time, there will not be enough at night when recombination of the ions has taken place in the lower layers of the atmosphere and it is necessary to deviate rays through larger angles to bring the waves down to the same receiver. Thus the outer radius of the silent zone should increase.

For longer distance transmission, where the bending is adequate and not electron-limited both by night and by day, we should expect an increase of signal intensity at night due to the reduced

Digitized by Google

absorption, so that the maximum range of audible signals should increase. This, again, is in agreement with the experimental facts.

If we assume that the sun is the source of ionization during the day the ionized layer should be lower in summer day-time than in winter day-time. Thus, according to the explanation given above, the outer radius of the silent zone and the maximum range of audibility should be greater in the latter case than in the former. Since the ionizing agent is removed at sunset, recombination takes place, so that the height of the layer responsible for ray deviation is, in general, higher during the night than during the day. The silent zone will thus have its maximum radius and the maximum long-distance ranges will therefore be attained at night in winter, when the layer is at its highest level during the year.

Since only one ray reaches a point just outside the silent zone it is clear that signal variations due to interference phenomena, such as have been shown to exist with broadcasting wave-lengths, cannot occur. Now it is remarkable that fading is less pronounced just outside the silent zone than at greater distances. The fading that exists at the shorter distances is thus most probably due to changes in the polarization of the waves. It may be noted that, according to the magneto-ionic theory of atmospheric influence, the plane of polarization of ultra-short waves is rotated by transmission through ionized gas in the earth's magnetic field whether the propagation is east and west or north and south.

Let us now consider the relation between the outer radius D of the silent zone (the "skipped distance" as it is sometimes called) and the wave-length. Let AB in Fig. 3 (which is not to scale) represent part of the circumference of the earth, the centre of which is at O.

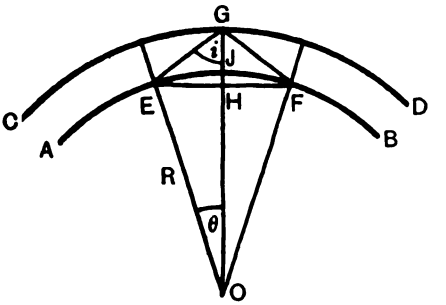


Fig. 3

Let CD similarly represent the ionized layer, so that JG the height of the layer is h. The wireless ray EG is incident at an angle i on the layer at G. EJF is the distance D of transmission and R

Digitized by Google

is the radius of the earth. If the angle EOH is θ it is easily shown that

$$D = 2R\theta \qquad \text{.....(6),}$$

and

$$\cos^2 i = \frac{(h + R - R\cos\theta)^2}{(h + R - R\cos\theta)^2 + R^2\sin^2\theta} \qquad \text{.....(7).}$$

But, in the cases that we are considering, θ is small, so that from (4) and (7) we have

$$\cos^2 i = \frac{4\pi Ne^2}{mp^2} = \frac{\left(h + \frac{R\theta^2}{2}\right)^2}{\left(h + \frac{R\theta^2}{2}\right)^2 + R^2\theta^2}.$$

On substituting for D from (6) this becomes

$$\frac{4\pi Ne^2}{mp^2} = \frac{\left(h + \frac{D^2}{8R}\right)^2}{\left(h + \frac{D^2}{8R}\right)^2 + \frac{D^2}{4}} \qquad \text{.....(8).}$$

There is therefore no very simple exact relation between p and D, but, since in most practical cases the first term of the denominator of the right-hand side of (8) is small compared with the second, we may write

$$\frac{4\pi Ne^2}{mp^2} = \frac{\left(h + \frac{D^2}{8R}\right)^2}{\frac{D^2}{4}} \qquad \text{.....(9),}$$

or

$$\frac{\lambda D}{h + \frac{D^2}{8R}} = \text{const.} \qquad \text{.....(10),}$$

λ being the wave-length.

Thus the product of λD should increase slowly with increasing values of D. What experimental evidence there is supports this.

Due to the earth's curvature there is, however, a limit to the value of D, when the deviation of the ray is a minimum. This case is obviously that in which the ray leaves the transmitter along the tangent to the earth at that point. In this case it is easy to show that the maximum value of D ($D_{\max.}$) is given by

$$D^2_{\max.} = 8Rh \qquad \text{.....(11),}$$

so that on substituting for R in (10) we find

$$\frac{\lambda D}{D^2_{\max.} + D^2} = \text{const.}$$

It is also of interest to consider the effects of electron-limitation in the case of maximum D and calculate the minimum wave-length

Digitized by Google

($\lambda_{\text{min.}}$) for which the minimum deviation referred to above can be brought about. On substituting the maximum value for D in (9) we find that

$$\frac{Ne^2}{\pi mc^2}\lambda^2_{\text{min.}} = \frac{2h}{R} \qquad \text{.....(12).}$$

Thus certain rays of the wave-lengths greater than the value given by (12) are imprisoned by the atmosphere and brought to the ground, but all rays of wave-length less than the critical value escape from the atmosphere altogether. If we again take N as 10^5 per c.c. and h as 60 km. we find that $\lambda_{\text{min.}}$ is about 10 metres. Too much importance should not be attached to the exact value of 10 metres in view of the uncertainty in the values of N and h, but it appears that wave-lengths less than a certain critical value in this range are of no use for long-distance communication in that the atmosphere does not bend them back to the ground. It is suggested that experiments to find the critical wave-length $\lambda_{\text{min.}}$ might be used to estimate N or h if one or the other is known.

In connection with the above discussion of imprisoned and escaping rays and wave-lengths it should be noted that, according to the magneto-ionic theory, there are cases in which, due to the influence of the earth's magnetic field, certain rays from long wave-length transmitters are split up into two elliptically polarized components for one of which the refractive index is greater than unity. This component is therefore not deviated back to the ground and so escapes via the atmosphere.

In the above calculations it has been assumed, for simplicity, that the deviation of the ray takes place at one point in the layer. This, of course, cannot be the case; the bending must be gradual. If a case of gradual bending of the rays be considered it is easy to show that the maximum value of N might be slightly larger than 10^5 electrons per c.c., but that it is unlikely that it is larger than 10^6 per c.c.

SUMMARY.

An interpretation of ultra-short wave wireless phenomena is given which indicates that the maximum number of electrons per c.c. in the atmospheric ionized layer is of the order 10^5 to 10^6.

In conclusion, I wish to express my indebtedness to Mr C. W. Goyder and Mr A. H. Cooper for experimental data on this subject.

Note added March 17th. After the proofs of this paper had been returned to the printer the February issue of the *Physical Review* reached London (March 15th), containing an article by A. Hoyt Taylor and E. O. Hulbert dealing with the same subject. Similar conclusions are arrived at regarding the maximum electronic content of the ionized layer.

Digitized by Google

The influence of an illuminated mercury surface on the Franck-Cario reactions. By HERBERT SIM HIRST, Trinity College. (Communicated by Mr E. K. RIDEAL.)

[*Received* and *read* 18 January 1926.]

Until the introduction of the modern quartz mercury vapour lamp the difficulty of obtaining an intense, and at the same time uniform source of light, rendered all investigations of the photochemical effects of ultraviolet radiation rather uncertain. In 1909, however, a series of communications began to appear by Berthelot and Gaudechon* describing a systematic examination of the effects of ultraviolet light in causing the combination of gases, such as hydrogen and oxygen, hydrogen and carbon monoxide, carbon monoxide and oxygen, etc. In some cases moderately large rates of reaction were observed, but the experiments were qualitative rather than quantitative. It is noteworthy, also, that in all these experiments the gases were enclosed in a quartz tube over a mercury surface. Since then a large amount of work has been done on this and similar subjects, but in recent years experimenters seem to have been more concerned with the effect of small traces of water vapour on the reaction in question than with the simple photochemical reactions themselves. It is not intended to enter into this part of the question, uncertain as it still is, and it must suffice to mention the work of Baker and Carlton† and of Coehn and Tramm‡.

In 1922 renewed interest was aroused by the publication by Franck and Cario§ of results announcing the production of an activated hydrogen by means of collisions of the second kind with mercury atoms excited to the $2p_2$ state by absorption of the resonance radiation, of wave-length 2536·7 Å. It was shown by the above authors that this active hydrogen was capable of reducing various metallic oxides, and it has also been established more recently by Dickinson|| and Mitchell¶ that it can also react directly with oxygen to form water. The precise nature of the activated state still seems uncertain. It was first claimed by the original discoverers that hydrogen atoms were produced, then Compton and Turner** postulated the existence of an unstable mercury hydride, while Mitchell†† favours the idea of an activated molecule of

* *Compt. Rend.* vol. 150, pp. 1169 *et seq.* (1910).
† *J. C. S.* vol. 127, p. 1990 (1925).
‡ *Ber.* vol. 54, p. 1148 (1921).
§ *Zeit. f. Physik*, vol. 11, p. 161 (1922).
|| *Proc. Nat. Acad. Sci.* vol. 10, p. 409 (1924).
¶ *Proc. Nat. Acad. Sci.* vol. 11, p. 458 (1925).
** *Phil. Mag.* vol. 48, p. 360 (1924).
†† *Loc. cit.*

Digitized by Google

hydrogen. The above experiments were all conducted at fairly low pressures, in the neighbourhood of half a millimetre of mercury, and the observed effects were small compared to those observed by Berthelot and Gaudechon.

More recently Taylor and Marshall* have carried out experiments at pressures of about half an atmosphere and have obtained truly enormous rates of reaction for the combination of hydrogen with oxygen, ethylene and carbon monoxide in the presence of mercury. The velocities are calculated to be at least some two hundred and fifty times greater than those of Dickinson. The largeness of the effect is ascribed to some form of chain mechanism, following the primary activation of the hydrogen by the excited mercury atoms. The present paper deals with a further experimental examination of the effects of ultraviolet light, strong in the mercury resonance line, $\lambda = 2536{\cdot}7$ Å. on the combination of hydrogen with various other gases in the presence of mercury and results of considerable interest have been obtained. Briefly it may be stated that the large reaction velocities of Taylor can only be obtained on illumination of a liquid mercury surface, and all experiments with mercury vapour alone have yielded but low values. This influence of an illuminated mercury surface on the reaction velocity seems to have been unsuspected hitherto, and its further examination promises to have considerable bearing on the work of previous experimenters.

Experimental Procedure. In the experiments to be described, the reaction chamber consisted of a transparent quartz tube of bore five-eighths of an inch, of which some 10 cms. in length could be exposed to the direct radiation from the lamp. The total volume of the reaction vessel and manometric system was calculated to be approximately 30 cubic centimetres. The quartz tube was connected to the rest of the apparatus, namely the gas inlet tubes and the manometer and pumps in some cases by sealing-wax joints and in the others by rubber stoppers, which were always well shielded from the ultraviolet rays. The whole system was so arranged that either a static or a dynamic flow method could be used at will. The gases were all passed slowly through soda lime, calcium chloride and phosphorus pentoxide tubes before admission to the reaction vessel, while the hydrogen and nitrogen were purified from oxygen by passage over hot copper turnings. The carbon monoxide was prepared from oxalic and sulphuric acids and the ethylene from alcohol and phosphoric acid.

The quartz mercury lamp was of the horizontal pattern, the centre tube of which was about 11 cms. in length, and the reaction vessel was placed parallel to it and at a distance of about 2 mm.

* *J. Phys. Chem.* vol. 29, p. 1140 (1925).

Digitized by Google

from it. The first attempts were made with the whole lamp immersed in a thermostat, as described by Taylor, but owing to the variations in the large current required and the very great condensation of mercury occurring on the walls, this method was abandoned. Finally the lamp was used as shown in the accompanying diagram (Fig. 1). Provided the water level was maintained constant a very steady light was obtained. The lamp was always used in conjunction with an electromagnet, in such a position that the discharge was pressed against the side of the quartz lamp, nearest to the reaction vessel. The lamp was run off 110 volts, the actual voltage drop across the lamp being some 21 volts, while the current was maintained as steady as possible at 2·6 amperes. Before every experiment, the lamp was strongly heated with a Bunsen flame as suggested by Baker. Under such conditions, a powerful source of ultraviolet radiation was obtained, which was at the same time fairly constant from experiment to experiment, in so far as could be judged from the power of reproducing results.

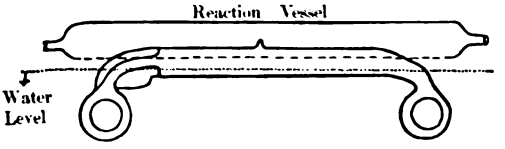


Fig. 1

Presentation of Results. The mercury lamp was run in all cases under the standard conditions described above. No attempt is made to preserve the original order in which the experiments were performed, and since the work has so far been mainly qualitative, the mean of numerous experiments is always presented.

Hydrogen and Oxygen. Although it is denied that a true photochemical union of the above gases does take place when they are exhaustively dried*, since in the present case neither reaction vessel nor gases were subjected to such extreme drying, it seemed necessary to perform several blank experiments. The back diffusion of mercury vapour from the manometer was prevented by a cadmium trap. Owing to the large heating effect of the mercury lamp, the temperature of the gases in the reaction vessel was considerably above room temperature, and the temperature quoted is the average temperature during the course of the reaction. Attempts were made to water jacket the reaction vessel, but with a thickness of water of 3 mm., all reaction stopped owing to the opacity of the water, and even with a thin film of water running over the tube the velocity was cut down to less than one-third.

* Baker, *loc. cit.*

Digitized by Google

It was found, however, that temperature equilibrium was soon established, and since the time for experiment was always long, the comparatively small temperature fluctuations were of no great importance. Table I shows the values obtained.

Table I.

Initial pressures in mm.		Total pressure	Temp. ° C.	Pressure decrease in mm. per hour	Conditions
H_2	O_2				
188·5	94·5	283	30	1·7	No mercury vapour present
163·5	81·0	244·5	30	1·65	Liquid mercury present completely shielded from light

It is noteworthy that the figures represent a very much smaller rate of reaction than those observed by Coehn and Tramm (10 per cent. change in five hours compared to 3 per cent. above) and Baker. It was also found that no measurable reaction took place at all unless the lamp was run under the conditions described, to ensure a radiation rich in the ultraviolet. It was such failures that first led us to examine what would be the result of illuminating the liquid mercury surface itself, and at once the larger reaction velocities were obtained, velocities approximately proportional to the area of mercury irradiated. In the experiments of Marshall and Taylor liquid water was always present in the reaction vessel, but this was not found necessary, indeed the reaction went with gases and vessel dry in the ordinary sense of the word, though whether it will go under Baker-dry conditions is quite another question. The following table presents a summary of the results. (Ten millimetres change of pressure corresponds to a disappearance of 0·356 c.c. of gas at N.T.P., i.e. to a disappearance of $9·6 \times 10^{18}$ molecules. Hence in the case of hydrogen-oxygen mixtures, twice this number of hydrogen molecules must react to cause a decrease of pressure of 10 millimetres.)

Table II.

Total pressure in mm.	Length of Hg surface exposed in cm.	Pressure decrease in mm. per hour	Temp. °C.
313	10·5	41	29
320	7·0	26	30
413	2·5	13·5	30
283	0·0	1·7	30

Digitized by Google

The gases were in all cases present in the ratio of two of hydrogen to one of oxygen. For the sake of comparison the velocities of reaction observed for the combination of hydrogen with ethylene to give ethane and with carbon monoxide to give formaldehyde are introduced here.

Table III.

Pressure H_2 in mm.	Pressure gas in mm.	Hg length in cm.	Pressure decrease in mm. per hour	Temp. ° C.	Remarks
198	117 CO	0	App. 1·0	31	—
198	117 CO	5·4	38·0	31	Reaction slows up as it proceeds, i.e. after 1 hour 26 mm. per hour
197	134 C_2H_4	0	4·6	28	—
194	131 C_2H_4	6·3	64·0	29	Steady velocity for long period

The retardation as the reaction proceeds in the case of the hydrogen-carbon monoxide mixtures is apparently due to a poisoning of the mercury surface by the liquid polymerisation products of the reaction. The greatest velocities were obtained with hydrogen-ethylene mixtures, and since this reaction proceeds smoothly, with no apparent detriment to the mercury surface, it would seem to be well suited for a more accurate study of the effect.

A very interesting side reaction was noted in the case of the combination of hydrogen and oxygen, namely the formation of a black film of mercuric oxide on the mercury surface and of a yellowish deposit of oxide on the inside of the quartz tube remote from the light. It was found on further investigation that this film was only formed in the presence of both hydrogen and oxygen. With pure oxygen traces of tailing of the mercury due to the formation of ozone by the ultraviolet rays of short wave-length (< 2000 Å.) were noticed, but there was no trace of the black film of mercuric oxide. The film is also only formed in the direct path of the light. With commercial hydrogen from a cylinder the blackening of the mercury takes place practically instantaneously, due to the presence of small quantities of oxygen as impurity. It must be concluded that the formation of this mercuric oxide is directly connected with the combination of hydrogen and oxygen and that a continuous removal of mercury vapour from the gas phase takes place.

The difficulty of illuminating a well-defined surface area of mercury in the liquid state led us to examine the efficiency of an amalgamated surface. With a sheet of carefully amalgamated silver whose area was 6·1 sq. cm. the combination of hydrogen and oxygen

Digitized by Google

proceeded at a rate corresponding to a pressure decrease of 60 mm. per hour. Although no definite explanation can yet be advanced, it may be of interest to mention that the blackening of the surface due to oxide formation occurred only on the back of the silver strip which was unexposed to the radiation. In the case, however, of a strip of amalgamated brass, of area 12·9 sq. cm., no reaction whatsoever could be detected, and it seems plausible to connect this with the presence of zinc in the surface layer and the consequent alteration in the vapour pressure of the mercury.

The negative results of Taylor in his attempts to cause the combination of carbon dioxide and hydrogen were corroborated, for with the manometric system in use, capable of measuring a pressure change of less than half a millimetre, no reaction could be detected. A still more remarkable result was obtained with mixtures of carbon monoxide and oxygen, which, according to Berthelot* and Coehn and Tramm†, combine to the extent of about 5 per cent. per hour under the influence of ultraviolet light. In the present experiments, although the light used was richer in the middle ultraviolet than in the case of the work of either of the above experimenters, practically no reaction was observed. Actually, a decrease in pressure of only half a millimetre per hour over a period of some five hours was noted, when the original gas pressure was 265 mm. of mercury, the two gases being present in the ratio of two of carbon monoxide to one of oxygen. An explanation of this apparently contradictory result still remains to be found.

Since the union of nitrogen and hydrogen possesses especial interest, it was decided to re-examine this reaction with great care, despite the fact that under Taylor's experimental conditions no combination had been detected. However the photosynthesis of ammonia from its elements in the presence of mercury had previously been effected by Noyes‡ at the temperature of boiling mercury and at pressures slightly greater than atmospheric, and it seemed probable that such a synthesis ought to take place, even at ordinary temperatures. A flow method was used in which the gases, mixed in known proportions, were slowly streamed at reduced pressure over a clean mercury surface, illuminated by the mercury lamp running under the conditions described above. Very great care was taken to remove all oxygen from the gases, since the presence of oxygen caused a rapid tarnishing of the mercury. The rate of flow of the gases was usually about one and a half litres per hour and the average pressure maintained in the reaction vessel was about 30 centimetres of mercury. After passage over the irradiated mercury surface, the gas mixture passed through a trap

* *Compt. Rend.* vol. 150, p. 1327 (1910).
† *Ber.* vol. 54, p. 1148 (1921).
‡ *J. Am. C. S.* vol. 47, p. 1003 (1925).

Digitized by Google

containing water and cooled to liquid air temperature. At the finish of a run this was sealed off and allowed to warm up, being well shaken all the while. The resultant solution was then subjected to careful tests with Nessler reagent and other test substances. Five experiments in all were conducted under the above conditions, the duration of a run varying from thirty minutes to three hours, while in addition two experiments were carried out at atmospheric pressure. In every case a very definite Nessler test for ammonia was obtained and there can be no doubt that ammonia *is* synthesised under the above conditions. This is all the more remarkable, since no light filter was available at the time to screen off the rays of wave-length shorter than 2200 Å. which are known to cause the dissociation of ammonia, so that it is probable that by far the largest part of the ammonia formed was destroyed before it passed out of the illuminated portion of the tube.

Similar experiments were next carried out with very dilute sulphuric acid in the liquid air trap since it was considered there might be a possibility of the formation of hydrazine and it was hoped to stabilise this as hydrazine sulphate. Extreme care was taken in the matter of blank tests, and it seems difficult to ascribe the positive results obtained in the tests to anything else than hydrazine, though sufficient was not isolated in any single case to give an absolutely distinctive test, such as the characteristic reaction with benzaldehyde. In every case, however, the solution obtained rapidly decolorised a dilute solution of potassium permanganate, and in two experiments traces of silver were precipitated from silver nitrate solution. That the decolorisation of the potassium permanganate was certainly not caused by traces of hydrogen peroxide was most conclusively shown by other tests with reagents such as starch-potassium iodide solution and titanous sulphate. Indeed, the view that hydrazine is formed is rendered all the more probable by analogy with the perfectly definite evidence of the formation of hydrogen peroxide in the case of hydrogen-oxygen gas mixtures. The experiments were conducted in a similar manner to those with nitrogen and hydrogen. The solution taken from the liquid air trap was in this case tested (*a*) with acid potassium permanganate solution, (*b*) with starch-potassium iodide solution in presence of ferrous sulphate, (*c*) with titanous sulphate, and in every test the characteristic reaction with hydrogen peroxide was observed, proving conclusively that it was the peroxide, and not traces of ozone, which was present.

It is known that hydrogen, activated by Wood's method in a high tension discharge, will react with sulphur to form hydrogen sulphide, which can then be detected with lead acetate solution. A similar experiment was tried in the present case by passing pure hydrogen over the irradiated mercury surface and then over

Digitized by Google

sulphur, but no trace of hydrogen sulphide was found. This shows either that the active hydrogen produced by impact with excited mercury atoms has a much shorter life than Wood's hydrogen, or that if it does persist long enough to reach the sulphur, it is not capable of reacting with it. In either case, it must be quite a different activated form of hydrogen to that produced by Wood's method.

Further experiments are now in progress, and it is hoped that with the apparatus now in use accurate quantitative data will be obtained. It seems rather remarkable that, considering the number of experiments previously performed on the photochemical union of the gases used in the present case, the phenomenon should never have been noted. At the present stage it is still rather difficult, indeed almost impossible, to draw any definite conclusions as to the precise mechanism of the phenomenon. It does seem, however, to be closely connected with the work of Franck and Cario and to depend on some primary activation of mercury atoms in close proximity to a mercury surface and the consequent activation of the hydrogen by the mercury. It is also clear that the concentration of active mercury atoms, during the period of illumination, is much higher close to the surface than in the gas phase. In this connection it seems of importance to note also the work of Bonhoeffer* on the chemiluminescence of active hydrogen. He found that with liquid mercury present, mercury hydride bands appear momentarily at the surface of the mercury in the spectrum of the active hy drogenand that this effect is solely heterogeneous in nature. The line $\lambda = 2536{\cdot}7$ Å. is also emitted under these conditions.

As regards the formation of the oxide film some such mechanism as the following may be postulated.

(*a*) Assuming the mercury hydride as an intermediate.

Hg′ represents an active mercury atom.

$$Hg' + H_2 \rightarrow Hg - H_2,$$

$$Hg - H_2 + O_2 \rightarrow H_2O_2 + Hg,$$

$$H_2O_2 + Hg \rightarrow HgO + H_2O,$$

and perhaps under the influence of the ultraviolet radiation,

$$HgO + H_2' \rightarrow Hg + H_2O.$$

The hydrogen peroxide will also undergo decomposition under the influence of the light.

* *Z. physikal. Chem.* vol. 116, p. 391 (1925).

Digitized by Google

Or (*b*) Assuming an activated hydrogen molecule as the intermediate.

$$H_2' + O_2 \rightarrow H_2O_2,$$
$$H_2O_2 + H_2' \rightarrow 2H_2O,$$
$$H_2O_2 + Hg \rightarrow H_2O + HgO,$$
$$HgO + H_2 \xrightarrow{\text{light}} Hg + H_2O.$$

At any rate in accordance with the experiments on the combination of nitrogen and hydrogen and of hydrogen and oxygen, it would seem that the syntheses of ammonia and water go by steps as follows:

(A) 1. $N_2 + H_2' \rightarrow N_2H_2$ (very unstable).

2. $N_2H_2 + H_2' \rightarrow N_2H_4$.

3. $N_2H_4 + H_2' \rightarrow 2NH_3$.

(B) 1. $H_2' + O_2 \rightarrow H_2O_2$.

2. $H_2O_2 + H_2' \rightarrow 2H_2O$.

In conclusion it may be said there still remains a very large quantity of work to be done before the phenomenon can be at all satisfactorily explained, and it does seem of considerable importance, since by far the greater part of the experimental work on the photochemical union of hydrogen with other gases has been done in the presence of mercury, which may or may not have influenced the results. If it has done so, the magnitude of the effect of the mercury must be determined before the results are of use.

Summary. It has been found that illumination of a mercury surface by ultraviolet light strong in the line $\lambda = 2536{\cdot}7$ Å., exerts a marked influence on the photochemical union of hydrogen and gases such as oxygen, ethylene, and carbon monoxide. The rate of reaction is proportional to the first order to the area of surface exposed.

The formation of a mercuric oxide film has been shown to occur only in the presence of a mixture of hydrogen and oxygen, and alternative mechanisms for its formation are suggested.

In the case of nitrogen-hydrogen mixtures, hydrazine and ammonia have been identified in the reaction products, and in the case of hydrogen and oxygen, hydrogen peroxide and water have been found, and it is suggested that the reactions proceed in steps.

The catalytic efficiency of the surface depends on its cleanness and is cut down by poisons, such as the reaction products in the case of the combination of hydrogen and carbon monoxide.

Digitized by Google

Since the communication of this paper, a letter from Dr Taylor and his co-workers has appeared in *Nature*, stating that high rates of reaction can be obtained with a flow method by bubbling the gases through mercury just before passage into the reaction chamber, and that our failure to obtain rapid reaction in a static system was due to the removal of the free mercury atoms from the illuminated zone. Our experiments lead us to the conclusion that this removal must be so rapid that the reaction is confined to a very small volume, immediately above the mercury surface, which acts as the source of a continuous supply of fresh mercury vapour. It is, of course, possible that mercury vapour may lose its photo-sensitising action after a certain time, though still present in the gas phase and that only newly-evaporated atoms arriving by diffusion possess this property.

In either case the ratio of the reaction velocities close to the surface and in the gas phase in a closed system will be large, and if the mercury vapour is readily subject to poisoning by the polymerisation products of side reactions, the surface is far less sensitive than the vapour.

Digitized by Google

The Behaviour of the Centres of Activity of Saturated Surfaces during the initial stages of Unimolecular Reactions. By F. H. CONSTABLE, B.A., Fellow of St John's College.

[*Read* 1 February 1926.]

The theory of catalytic action recently advanced qualitatively by Pease*, Taylor†, Armstrong and Hilditch‡, Constable§, which the author has endeavoured to treat quantitatively||, is extended in this paper to the explanation of the effect of pressure and inert diluents on catalytic action on saturated surfaces. The results are shown to be in agreement with experiment¶, and further inferences may be made concerning the relative mean lives of the reactant and the diluent on the surface.

There is a great deal of experimental material which throws light on the state of the adsorbed molecules of the centres of activity. Benton** following Langmuir†† distinguishes two types of adsorption; the primary adsorption, which increases very rapidly with the pressure at low pressures, but soon reaches a saturation value beyond which the quantity adsorbed remains independent of further pressure increase; and secondary adsorption in which the amount of gas adsorbed increases till really high pressures are obtained. The two types of adsorption usually occur together. From a study of the catalytic oxidation and adsorption of carbon monoxide Benton finds that the primary adsorption of carbon monoxide on a series of metallic oxides is in the same order as the catalytic activity.

Pease*, as a result of much experimental work, shows that it is the strong low pressure adsorption which is responsible for catalytic activity. From observations cited by Taylor† it would seem that at high temperatures and low pressures the heat of adsorption may reach enormous values in some cases. From these observations it is deduced that the active centres are always covered with the reactant except at low pressures and high temperatures. Langmuir‡‡ developed a theory of adsorption which has

* *J. Amer. Chem. Soc.* vol. 45, pp. 1196, 2235, 2297 (1923).

† *Proc. Roy. Soc.* A, vol. 108, p. 105 (1925); also *J. Phys. Chem.* vol. 28, p. 911.

‡ *Trans. Farad. Soc.* vol. 17, p. 670 (1922); *Proc. Roy. Soc.* vol. 108, p. 111.

§ *Proc. Roy. Soc.* vol. 107, p. 274 (1925).

|| *Proc. Roy. Soc.* vol. 108, p. 374 (1925); *Nature*, vol. 116, pp. 278–9; vol. 117, pp. 230–237 (1926).

¶ See following papers.

** *J. Amer. Chem. Soc.* vol. 45, pp. 887, 900 (1923).

†† *J. Amer. Chem. Soc.* vol. 40, p. 1400 (1918); also vol. 38, p. 2267 (1916); and *Trans. Farad. Soc.* vol. 17, p. 606 (1922).

‡‡ *J. Amer. Chem. Soc.* vol. 38, p. 2267 (1916) *et seq.* (*loc. cit.*).

Digitized by Google

been considerably extended by Frenkel*, and it is proposed to develop these considerations still further in accord with the recent development of the theory of centres of activity. Experiment has shown that chemical change only occurs on active centres occupying a fraction of the surface, which are always covered with the reactant except at low pressures and high temperatures. In fact the time during which these centres are left free of adsorbed reactant is negligible compared with the time they are covered. The time during which an adsorbed molecule remains on the centre is practically independent of the rate of bombardment of the saturated surface, its release being caused by its energy of oscillation perpendicular to the surface exceeding the work of desorption; or in the case of chemical action by the reacting molecules receiving the critical increment in the required degree or degrees of freedom†. In either case as soon as the adsorbed molecule or its reaction products have evaporated, the next molecule bombarding the centre will remain‡. Thus over a wide pressure range the state on the active centres remains the same in the initial stages of the reaction, and consequently the rate of unimolecular chemical change at the surface is unaltered§.

With multi-molecular changes, and when the partial pressure of the single reactant is diminished by diluents, many types of molecule may be adsorbed by an active centre. Thus when one molecule has evaporated it will be replaced by the next one bombarding the surface, which may be either reactant or diluent. Thus the time during which an active centre is covered by the reactant is reduced. Hence, although reducing the pressure has no effect on the velocity of unimolecular decomposition on saturated surfaces over a considerable range, the presence of a diluent may decrease the reaction velocity, the amount of the decrease depending on the relative mean lives of the reactant and diluent.

There is a considerable amount of evidence supporting this deduction, which was first made by Bodenstein and Fink‖ who observed that the presence of sulphur trioxide retarded the oxidation of sulphur dioxide. Carbon dioxide retards the oxidation of carbon monoxide¶. Water retards the catalytic dehydration of alcohol** and ether††, and the two decompositions of formic acid

* *Zeit. für Phys.* vol. 26, p. 117 (1924).

† For experimental evidence see Norrish, *J. Chem. Soc.* vol. 123, p. 3007; Hinshelwood and Pritchard, *J. Chem. Soc.* vol. 123, p. 2725 (1923).

‡ This is a fundamental hypothesis of the Langmuir-Frenkel theory.

§ There is very strong evidence in support of this constancy for alcohol and copper; it appears however that at very high pressures the velocity should fall with pressure, because the escape of the molecule is retarded.

‖ *Zeit. Phys. Chem.* vol. 60, p. 61 (1907); Bodlander and Kopper, *Zeit. Electrochem.* vol. 9, p. 556 (1903); Berl, *Zeit. anorg. Chem.* vol. 44, p. 267 (1905).

¶ Henry, *Phil. Mag.* (iii), vol. 9, p. 324 (1836).

** Engelder, *J. Phys. Chem.* vol. 21, p. 676 (1917).

†† Ipatieff, *Ber. deutsch. chem. Ges.* vol. 37, p. 2996 (1904).

Digitized by Google

on alumina*. Lamb and Vail† have also made a complete series of measurements on the retarding action of water on the oxidation of carbon monoxide, and hydrogen on the surface of a mixture of manganese dioxide and cupric oxide known as "Hopcalite". Larsen and Tour‡ report the retarding influence of water in the synthesis of ammonia. Thus the qualitative effect of a diluent has been frequently observed, and it is of interest to develop the quantitative treatment.

The Effect of Diluents on the Initial Stages of Unimolecular Surface Reaction.

Let p = the fractional pressure of the reactant in the mixture,

P = the total gas pressure,

m = the number of molecules which bombard 1 sq. cm. of the catalyst per second,

s = the area of a centre of activity,

and M = the molecular weight.

Then
$$m = \left(\frac{1}{2\pi MRT}\right)^{\frac{1}{2}} Pp = \mu p,$$
defining μ.

The rate at which the reactant molecules strike a centre is
$$\mu_1 sp.$$
The rate at which the diluent molecules strike is
$$\mu_2 s\,(1-p).$$
Thus the probability that a reactant molecule should strike a given centre at any instant is
$$\frac{\mu_1 p}{\mu_1 p + \mu_2 (1-p)}.$$
When a molecule or its decomposition products have evaporated, the next molecule bombarding the centre is adsorbed. Thus if in some time interval t, the centre is left vacant n times,
$$\frac{\mu_1 pn}{\mu_1 p + \mu_2 (1-p)}$$
molecules of the reactant, and
$$\frac{\mu_2 (1-p)\, n}{\mu_1 p + \mu_2 (1-p)}$$
molecules of the diluent will condense.

* Adkins and Nissen, *J. Amer. Chem. Soc.* vol. 45, p. 809 (1923).
† *J. Amer. Chem. Soc.* vol. 47, p. 123 (1925).
‡ *Chem. Met. Eng.* vol. 26, p. 650 (1922).

Digitized by Google

If τ be the mean life of a molecule on a surface,

τ_0 be the vibration frequency perpendicular to the surface,

and μ_0 be the heat of desorption per gram mol,

then Frenkel has shown that

$$\tau = \tau_0 e^{\mu_0/RT}.$$

Let τ' and τ'' be the mean lives of reactant and diluent on active centres (defined by the heat of activation ϵ) as in the preceding equation. Then the centre is occupied by the reactant for a time

$$\frac{\mu_1 \tau' pn}{\mu_1 p + \mu_2 (1-p)},$$

and by the diluent for a time

$$\frac{\mu_2 \tau'' (1-p)\, n}{\mu_1 p + \mu_2 (1-p)}.$$

Since the time the surface is bare is negligible compared with the time it is covered, the fraction of the total time a centre is occupied by the reactant is

$$\frac{\mu_1 \tau' p}{\mu_1 \tau' p + \mu_2 \tau'' (1-p)} = \frac{1}{1 + \lambda (1/p - 1)},$$

where
$$\lambda = \frac{\mu_2 \tau''}{\mu_1 \tau'} = \frac{\mu_2 \tau_0'' e^{\mu_0''/RT}}{\mu_2 \tau_0' e^{\mu_0'/RT}} = \frac{\tau_0'' e^{\mu_0''/RT}}{\tau_0' e^{\mu_0'/RT}} \left(\frac{M_1}{M_2}\right)^{\frac{1}{2}}.$$

Thus though λ is independent of p, it is not independent of T unless the heats of desorption of the two molecules are equal.

Let the distribution of the active centres over the catalyst be defined by the relation*

$$\delta n = F(\epsilon)\, \delta\epsilon$$

between the limits ϵ_1 and ϵ_2. Then the fractional reduction of the reaction velocity is the fraction of the total time the centres are occupied by the reactant, multiplied by the rate of reaction on each $f(\epsilon)$, integrated over the whole distribution.

Let v = the fractional reaction velocity at the fractional pressure of reactant p,

then
$$v = \frac{\int_{\epsilon_1}^{\epsilon_2} \frac{f(\epsilon)\, F(\epsilon)\, \delta\epsilon}{1 + \lambda (1/p - 1)}}{\int_{\epsilon_1}^{\epsilon_2} f(\epsilon)\, F(\epsilon)\, \delta\epsilon} \qquad \text{.....(1).}$$

Case I. If the reactant and the diluent are of the same chemical nature, then it would be expected that both types of molecules

* Constable, "Mechanism of Catalytic Decomposition", *loc. cit.*

Digitized by Google

would behave similarly as the surface fields on the reaction centres increase, and the quantity

$$\lambda = \frac{\mu_2 \tau_0'' e^{\mu_0''/RT}}{\mu_1 \tau_0' e^{\mu_0'/RT}}$$

would remain constant. It is also of interest to note that if the great majority of chemical action takes place on centres of the same type* (those centres with the strongest fields of the required nature) then the majority of the fall in reaction velocity would follow the law obtained by assuming λ a constant. Thus the formula obtained would be expected to be of a very general nature in spite of the fact that the ratio of the mean lives of the reactant and diluent has been assumed independent of the fields on the adsorbing centres.

Hence $$v = \frac{\dfrac{1}{1 + \lambda(1/p - 1)} \displaystyle\int_{\epsilon_1}^{\epsilon_2} f(\epsilon)\, F(\epsilon)\, \delta\epsilon}{\displaystyle\int_{\epsilon_1}^{\epsilon_2} f(\epsilon)\, F(\epsilon)\, \delta\epsilon},$$

or $$1/v = (1 - \lambda) + \lambda/p \qquad \text{.....(2).}$$

A formula containing one constant connects the fractional reaction velocity v with the fractional pressure of the reactant p. The constant λ depends on the ratio of the molecular weights and the mean lives of the two molecules on the surface. If the mean life of the diluent is very short λ is small, and very little reduction occurs in the reaction velocity. If $\lambda = 1$, the two types of molecules behave similarly, except that the one can react and the other cannot, and the formula becomes $v = p$, that is the reaction velocity is proportional to the partial pressure of the reactant. This is to be expected, hence the formula is confirmed. The case in which λ becomes large is also considered separately at the end of the paper, and provides explanation of the action of typical catalyst poisons.

The general forms of the curves of the equation are shown in Fig. 1. It is to be observed that the curve possesses a definite slope at the origin, and that this slope continuously changes till it reaches another definite value. The gradient of the curve is given by the equation

$$\frac{dv}{dp} = \frac{\lambda}{[(1 - \lambda)\, p + \lambda]^2}.$$

Thus when $p \to 0$, $dv/dp \to 1/\lambda$, when $p \to 1$, $dv/dp \to \lambda$, giving the gradient of the two tangents, and the value at $p = 1/2$ is $\lambda/(1 + \lambda)^2$, intermediate between the two tangents. In curve I ($\lambda = \cdot 01$) and very slight change is produced in the velocity except

* I.e. the most active in the distribution: see Constable, *Proc. Roy. Soc.* A, vol. 107, p. 376 (1925).

Digitized by Google

at small partial pressures; curve II shows ($\lambda = \cdot33$); curve III ($\lambda = \cdot66$); curve IV ($\lambda = 1$) and is a straight line; curve V ($\lambda = 2$); and curve VI ($\lambda = 100$) and is typical of a catalyst poison.

The effect of water on the initial stages of the dehydrogenation of alcohol by copper is a case of this formula in which $\lambda = 0{\cdot}36$*.

Case II. λ varies with the nature of the active centre. μ_1 and μ_2 are constants, the variation is therefore in the ratio of the mean lives of the molecules on the surface. The work of desorption of the two molecules increases as the field on the centres of activity

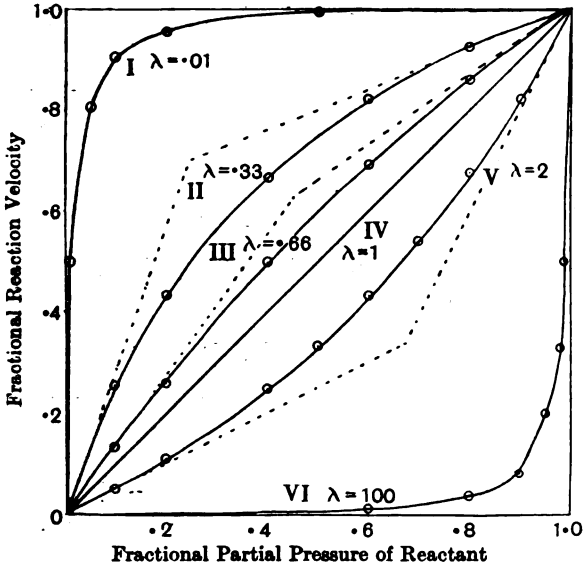


Fig. 1. Inert Diluents.

increases. When the work of desorption becomes very large the reactant molecule may be released much faster by receiving the critical increment in the appropriate degree of freedom, than by receiving the necessary energy of oscillation perpendicular to the surface, since the resultants are in general much less strongly adsorbed than the reactants. The diluent molecule, however, can only be released by virtue of its excessive amplitude of oscillation perpendicular to the surface. Thus for large fields μ_0' becomes constant while μ_0'' still increases with increasing fields, or in terms of the theory of activation previously advanced when the heat of activation falls to some critical value, λ increases rapidly.

* Constable, *Nature*, vol. 116, p. 739 (1925).

Digitized by Google

Let $\lambda = \lambda_0 \phi(\epsilon)$, then equation (1) becomes

$$v = \frac{\int_{\epsilon_1}^{\epsilon_2} \frac{f(\epsilon) F(\epsilon) \delta\epsilon}{1 + \lambda_0 \phi(\epsilon)(1/p - 1)}}{\int_{\epsilon_1}^{\epsilon_2} f(\epsilon) F(\epsilon) \delta\epsilon} \qquad \ldots\ldots(3),$$

and there is insufficient evidence to evaluate the case further. The equation would take the form

$$1/v = (1 - \lambda_0 F(p)) + \lambda_0 F(p)/p,$$

and the observations of Pease and Stewart* confirm that this case exists.

Bimolecular Reaction. The reactants must be adsorbed on the appropriate adjacent centres. The probability that the one centre will have reactant adsorbed is

$$\frac{1}{1 + \lambda_1 (1/p - 1)},$$

the probability that the second centre will have an adsorbed reactant molecule is

$$\frac{1}{1 + \lambda_2 (1/p - 1)}.$$

Thus the fraction of the total time both centres are occupied together is

$$\left(\frac{1}{1 + \lambda_1 (1/p - 1)}\right)\left(\frac{1}{1 + \lambda_2 (1/p - 1)}\right),$$

and this is the fractional reduction in reaction velocity.

Thus the equation becomes

$$1/v = (1 - \lambda_1 + \lambda_1/p)(1 - \lambda_2 + \lambda_2/p).$$

The case of λ variable with ϵ is too complicated to be considered at present. In some cases, however, owing to the ease with which the second reactant is adsorbed and activated, bimolecular reactions on surfaces appear unimolecular†, and the equations deduced previously apply. This will also be the case if there is a marked difference between λ_1 and λ_2.

Multimolecular Reactions. The general equation takes the form

$$1/v = \prod_{1 \to n} (1 - \lambda + \lambda/p).$$

* *J. Amer. Chem. Soc.* vol. 47, p. 1235 (1925).
† E.g. C_2H_4 and H_2 (Pease, *loc. cit.*).

Digitized by Google

The Composition of the Adsorbed Gas Film in contact with Mixed Gases and Vapours.

Let p_1, p_2, and p_n be the partial pressures of the gases in the mixture. Then with the same notation as before, the probability that a molecule of the first sort shall strike a centre of adsorption is

$$\frac{\mu_1 p_1}{\Sigma \mu_{1n} p_{1n}}.$$

Thus the fraction of the total time the centre is occupied with the first molecular species is

$$\frac{\mu_1 \tau' p_1}{\Sigma \mu_n \tau^{(n)} p_n}.$$

Considering the whole surface as homogeneous, the fraction of the whole surface covered by the species of pressure p_1 is also

$$\frac{\mu_1 \tau' p_1}{\Sigma \mu_n \tau^{(n)} p_n} = \frac{1}{1 + \lambda_1 p_2/p_1 + \lambda_2 p_3/p_1 + \lambda_3 p_4/p_1 \cdots \lambda_{n-1} p_n/p_1}.$$

Thus the fraction covered depends on the nature of each gas as well as its partial pressure in the mixture, and also gives the reduction in reaction velocity when many gases and vapours are present as diluents.

The Effect of Poisons on the Initial Stages of Unimolecular Surface Action.

When the mean life of the diluent on the surface is large a very small pressure of the diluent will reduce the reaction velocity to a negligibly small fraction of its initial value, and the diluent is called a catalyst poison.

Let $x =$ the fractional partial pressure of the poison. Then equation (2) becomes

$$v = \frac{1}{1 + \lambda x/(1 - x)}$$

$$= \frac{1}{(1 + \lambda x)},$$

since x is small*. Hence

$$1/v = (1 + \lambda x) = 1 + \left\{ \left(\frac{M'}{M''} \right)^{\frac{1}{2}} \left(\frac{\tau_0''}{\tau_0'} \right) \frac{e^{\mu_0''/RT}}{e^{\mu_0'/RT}} \right\} x,$$

the poisoning of the most active centres only being considered, λ being assumed constant. Thus the efficiency of a substance as a catalyst poison increases with increase of the period of oscil-

* Verified by Hinshelwood and Pritchard for poisonous action of O_2 on the decomposition of N_2O, *J.C.S.* vol. 127, p. 330 (1925): I have confirmed this poisonous action of O_2 on the platinum catalyst and results agree with the equation deduced. Under conditions to be discussed later, this formula is rigorously true, and no longer an approximation when the pressure of poison is small.

Digitized by Google

lation perpendicular to the surface of the catalyst and the heat of desorption, but decreases with the molecular weight. If the heat of desorption of the diluent is considerably different from the heat of desorption or the heat of activation of the reactant, or the sum of both, as the case may be, then the efficiency of the diluent as a poison will be considerably influenced by the temperature. In general when the diluent is a strong poison, its heat of desorption is much greater than that of the reactant, hence the higher the temperature the less pronounced the poisoning effect*. This effect of temperature is well known in manufacturing processes.

Continuous exhaustion of the catalyst should be sufficient to remove the poison completely. This has also received experimental confirmation†.

Bimolecular Reaction. The fraction of the total time during which two centres are occupied simultaneously becomes

$$v = \left(\frac{1}{1+\lambda_1 x}\right)\left(\frac{1}{1+\lambda_2 x}\right).$$

Hence $$1/v = 1 + ax + bx^2,$$

where $a = \lambda_1 + \lambda_2$, and $b = \lambda_1\lambda_2$, λ_1 and λ_2 being considered as constants. Both constants can be calculated from a knowledge of a and b. Thus

$$\lambda_1 = \tfrac{1}{2}\{a + (a^2 - 4b)^{\frac{1}{2}}\},$$

and $$\lambda_2 = \tfrac{1}{2}\{a - (a^2 - 4b)^{\frac{1}{2}}\}.$$

If b is small compared with a, then λ_2 is negligible, and the equation reduces to the form for unimolecular reaction. The observations of Pease and Stewart on the effect of carbon monoxide as a poison in the reaction of ethylene and hydrogen on copper will be discussed in following communications. At low temperatures the reaction velocity appears proportional to the pressure of hydrogen.

The case of λ_1 and λ_2 variable with the nature of the centre of activity is supported by the observations of Pease and Stewart, but is so far too complicated to work out without more experimental evidence.

It is interesting to notice that in the case of the poisoning of a copper catalyst by mercury, the order of the reduction of the reaction velocity by poisoning (1/200th) is of the order of the square of the reduction in the adsorption of hydrogen (1/20th),

* Hinshelwood and Burk, *J. Chem. Soc.* vol. 127, pp. 1114–6. The effect of hydrogen on the decomposition of NH_3 by platinum decreases as temperature rises. Langmuir, *Trans. Farad. Soc.* vol. 17, p. 621 (1922), shows that the poisonous action of H_2 and CO on their respective reactions to form H_2O or CO_2 is marked at low temperatures but vanishes as the temperature is raised.

† Pease, *loc. cit.*

Digitized by Google

the hydrogen gas being held chiefly by primary adsorption. This evidence is in favour of the method of treatment given above, and supports the obvious fact that the two reacting molecules must be present together before chemical reaction will take place.

The types of curves obtained from both laws are shown in Fig. 2, for the values of $\lambda = 10, 100, 1000$. The figure bears striking resemblance to those obtained by Lamb and Vail*. More experimental evidence in favour of the relations deduced here will be given in succeeding publications.

My thanks are due to the department of Scientific and Industrial Research for a grant in aid of this work.

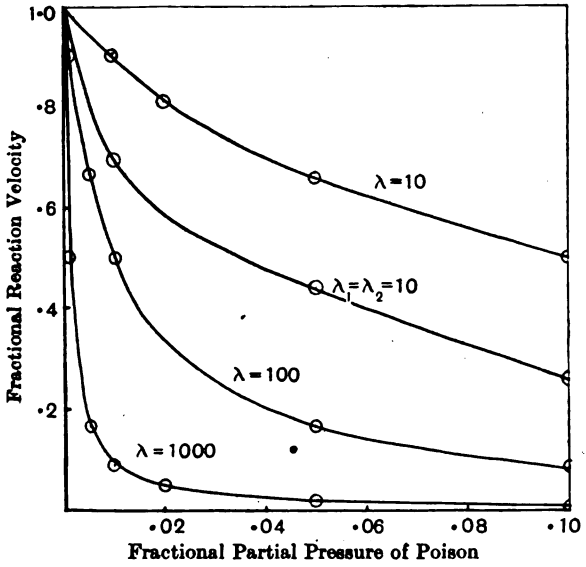


Fig. 2. Poisons.

Summary.

A considerable amount of evidence is given that while decrease in pressure of the reactant has little effect on reaction velocity of unimolecular reaction on saturated surfaces, inert diluents cause a marked decrease. An extension of the Langmuir-Frenkel theory to the centres of activity on the catalyst accounts satisfactorily for the observations, and quantitative relations are worked out independent of the mechanism of the chemical change.

* *J. Amer. Chem. Soc.* vol. 47, pp. 132, 133, figs. 4 and 5 (1925). The modification in this law of poisoning caused by the formation of a complex between the reactant and the poison which is more highly reactive than the adsorbed reactant alone will be considered in a later publication. When this occurs the curves show a point of inflection, as is seen in the observations quoted.

Digitized by Google

If p is the fractional pressure of the reactant in the mixture, and v is the fractional reaction velocity, then

$$1/v = (1 - \lambda) + \lambda/p,$$

where the constant λ depends on the ratio of the molecular weights, the periods of oscillation of the perpendicular to the surface, and the heats of desorption of the two molecular species. Thus though λ is independent of the pressure for homogeneous centres, it is only independent of the temperature if the heats of desorption of the two molecules are equal.

It is expected that cases would be found in which λ varies with the field on the reaction centre, when λ appears a function of the pressure.

The expression is extended for multimolecular reactions. These considerations also enable the composition of the adsorbed film in contact with a given gaseous atmosphere to be worked out.

An explanation is provided of many of the characteristic properties of catalyst poisons.

If x is the pressure of the poison, then

$$1/v = 1 + \lambda x$$

for unimolecular reaction, and

$$1/v = 1 + ax + bx^2$$

for bimolecular reaction. Since strong poisons in general have a high heat of adsorption, these expressions give explanation of the fact that the higher the temperature the less efficient the poison. Experimental evidence is shown to be in favour of the expressions deduced.

Digitized by Google

Metallic Complexes with the Aliphatic Poly-Amines. By Dr F. G. MANN and Prof. Sir WILLIAM J. POPE.

[*Read* 1 February 1926.]

The existence of complex compounds of metallic salts with ammonia was recognised in the early years of last century. Platinum in particular readily unites with ammonia, and the early chemists, when studying the general chemistry of ammonium platino- and platini-chlorides, soon discovered highly crystalline ammonia derivatives of platinum which contained an unusually high proportion of the metal. In the formulation of such compounds difficulties at once arose. Thus Peyrone in 1844 knew three distinct compounds of the composition $PtCl_2N_2H_6$, a pale lemon-coloured compound known as the second chloride of Reiset, an orange-coloured compound discovered by Peyrone himself, and finally a dark green insoluble substance known as Magnus' green salt, discovered by the latter in 1828. The existence of these three compounds, now known as trans and cis dichloro-diamino-platinum $[Pt(NH_3)_2Cl_2]$, and as tetramino-platinous platinochloride $[Pt(NH_3)_4]PtCl_4$ respectively, could not be satisfactorily explained on current theories of salt formation. Similar compounds in which aliphatic and cyclic mono-amines replaced the ammonia groups were discovered later, but it was not until 1889 that Jörgensen* introduced the use of the simplest stable aliphatic diamine, viz. ethylene diamine. With this amine he prepared the compounds now known as *bis*-ethylenediamine-platinous dichloride

$$[Pt(NH_2CH_2CH_2NH_2)_2]Cl_2,$$

and tri-ethylenediamine cobaltic trichloride

$$[Co(NH_2CH_2CH_2NH_2)_3]Cl_3.$$

Although the development of our knowledge of the stereochemistry of metallic complex compounds in the hands of Werner and his co-workers has been based almost solely on the use of ethylenediamine and other aliphatic diamines, and of similar divalent acidic groups, such as the oxalate and carbonate radicles, yet no further advance in the use of higher polyamines was made until 1924. A method of preparing $\alpha\beta\gamma$-triamino-propane, $H_2N\,CH(CH_2NH_2)_2$, in quantity was then worked out by the present authors†, and the triamine was used further to prepare complex compounds with the heavy

* *J. pr. Ch.* vol. 39, [2], p. 3 (1889).

† Mann and Pope, *Proc. Roy. Soc.* A, vol. 107, p. 80 (1924).

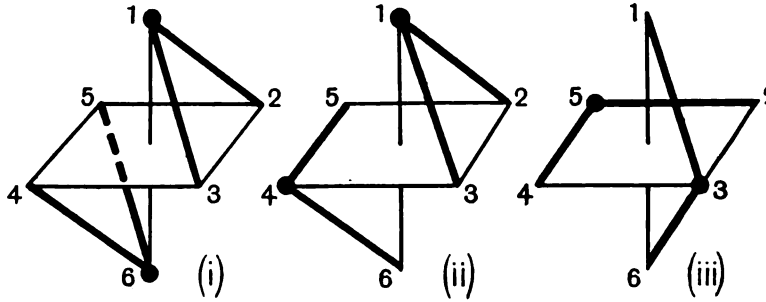
Digitized by Google

metals. With trivalent cobalt and rhodium, compounds of the general type of *bis*-propanetriamine-cobaltic trichloride,

$$[Co(NH_2CH(CH_2NH_2)_2)_2]Cl_3,$$

were prepared, and later similar compounds with divalent nickel were isolated. These compounds differ from Jörgensen's salts in that two molecules of the triamine replace three molecules of the diamine.

If Werner's octahedral configuration of metallic complexes of co-ordination number six be accepted, the above cobalt compound should exist in three isomeric forms: in these forms the two molecules of the triamine on the conventional system of notation occupy the following positions around the octahedron: (i) 2, 1, 3 : 4, 5, 6, (ii) 2, 1, 3 : 5, 4, 6, and (iii) 1, 3, 6 : 4, 5, 2.



The complex of type (i) possesses a plane of symmetry and will not be resolvable into optically active forms: type (ii) however should be so resolvable, whilst the complex of type (iii), if the two molecules of the triamine are symmetrically disposed about two cube planes of the octahedron, should not be resolvable. It was found in practice, however, that the complex compounds with cobalt, rhodium, and nickel were all highly crystalline compounds showing no trace of such isomerism, a result due presumably to the fact that the complex of type (i) is probably far more stable than those of types (ii) and (iii), and therefore alone occurs.

Attempts were therefore made to realise such isomerism by the use of other triamines, and recourse was had to a compound of completely different character, viz. $\beta\beta'\beta''$-triaminotriethylamine, $N(C_2H_4NH_2)_3$. This base had been prepared by Ristenpart*, who showed that in this compound the close proximity of the primary amine groups to the tertiary nitrogen atom apparently rendered the latter neutral. The base therefore gave a trihydrochloride, $N(C_2H_4NH_2)_3$, 3HCl, and acted as a triamine in all the derivatives which Ristenpart prepared. Since moreover no coordination com-

* *Ber.* vol. 29, p. 2530 (1896).

Digitized by Google

pound of an aliphatic tertiary amine had been recorded, it was expected that the base would give with cobalt a compound of the type *bis*-triaminotriethylamine-cobaltic trichloride,

$$[Co(N(C_2H_4NH_2)_3)_2]Cl_3.$$

Investigation showed, however, that the base can act as a weak tetramine, giving with strong hydrochloric acid an unstable tetrahydrochloride, $N(C_2H_4NH_2)_3$, $4HCl$, H_2O, and a number of simple derivatives in which the amine thus acts as a tetramine were described*. It was found, moreover, that the base coordinated with metals always as a tetramine, and gave with tetravalent platinum of coordination number six the compound dichlorotriaminotriethylamine-platinic dichloride, $[Cl_2PtN(C_2H_4NH_2)_3]Cl_2$. This dichloride is the first complex compound containing an aliphatic tetramine to have been prepared: it is moreover the first recorded compound in which an aliphatic tertiary nitrogen atom occupies a coordination position. With divalent platinum and divalent palladium, both of coordination number four, the tetramine gives compounds of the type of triaminotriethylamine platinous di-iodide, $[PtN(C_2H_4NH_2)_3]I_2$. Since no trace of a coordination compound in which the base acted as a triamine could be discovered, the question of the isomerism of such compounds did not arise.

In the case of the above compounds of divalent platinum and palladium, in which the metal shows a coordination number of four, it was considered that less strain upon the tetramine molecule would result if two molecules of the base shared two atoms of platinum, giving a compound of formula $[Pt_2(N(C_2H_4NH_2)_3)_2]I_4$. The compounds of platinum and palladium, being almost insoluble in cold water, were however ill suited to molecular weight determinations, and an investigation of the nickel complex salts was therefore undertaken. Divalent nickel may show a coordination number of four or of six, giving thus tetrammino-nickelous sulphate (iv) and hexammino-nickelous di-iodide (v). With triaminotriethylamine, complex salts of similar type were prepared, viz. triaminotriethylamine nickelous sulphate (vi) in which the metal has

(iv) $[Ni(NH_3)_4]SO_4$, (v) $[Ni(NH_3)_6]I_2$,

(vi) $[Ni\,N(C_2H_4NH_2)_3]SO_4$, (vii) $[Ni_2(N(C_2H_4NH_2)_3)_3]I_4$.

a coordination number of four, and *tris*-triaminotriethylamine-*bis*-nickelous tetraiodide (vii), in which the metal has a coordination number of six, three molecules of the tetramine sharing two atoms of nickel for this purpose. Triaminotriethylamine nickelous sulphate (vi) is thus of the same general type as the above compounds of divalent platinum and palladium, and being freely

† Mann and Pope, *Proc. Roy. Soc.* A, vol. 109, p. 444 (1925).

Digitized by Google

soluble in water lent itself to molecular weight determinations. The monosulphate of formula

$$[\mathrm{Ni\ N(C_2H_4NH_2)_3}]\mathrm{SO_4}$$

has a molecular weight of 300·9, and in dilute solution if ionisation is complete an "apparent" molecular weight of 150·4, whereas the disulphate of formula $[\mathrm{Ni_2(N(C_2H_4NH_2)_3)_2}](\mathrm{SO_4})_2$ has a molecular weight of 601·8 and in solution an "apparent" molecular weight of 200·6. The actual results obtained by the cryoscopic method in water are:

Concentration (grams anhydrous sulphate per 100 grams water)	"Apparent" molecular weight
·422	160 (157*)
1·32	173
1·79	180
3·23	218

These results provide strong evidence that the sulphate has the simple formula [Ni tren]$\mathrm{SO_4}$, and that the "apparent" molecular weight approaches 150 at great dilution, increasing as the concentration increases and the degree of ionisation falls. The fact that the equivalent conductivity of the sulphate at various dilutions approaches very closely to that of nickel sulphate itself, $\mathrm{NiSO_4}$, confirms the simple formula.

The nickel compound of coordination number six (vii) represents a novel type of di-complex salt, in which two ethylene residues act as bridging groups between the two nickel octahedra. It is difficult to see how such a complex can be other than dissymmetric, yet all attempts to resolve the tetraiodide failed. This result is in harmony with previous experience, since no record has yet appeared of the resolution of a complex nickel compound.

A second aliphatic tetramine, $\gamma\gamma'\gamma''$-triaminotripropylamine, $\mathrm{N(C_3H_6NH_2)_3}$, has also been prepared. In this substance, the tertiary nitrogen atom, by virtue of its greater distance from the primary amine groups, is strongly basic, and the compound, in both its simple and its complex compounds, acts therefore solely as a tetramine. With nickel, it gives coordination derivatives of the type of triaminotripropylamine nickelous dithiocyanate, $[\mathrm{Ni\ N(C_3H_6NH_2)_3}](\mathrm{SCN})_2$, in which the metal shows a coordination number of four. No complex compounds, either of nickel or of platinum, in which the metal shows a coordination number of six, could however be isolated.

* Calculated value, correct to only two significant figures.

Digitized by Google

Oxidations on Charcoal. By Miss W. M. WRIGHT. (Communicated by Mr E. K. RIDEAL.)

[*Read* 1 February 1926.]

It is well known that charcoal possesses catalytic properties. It accelerates the decomposition of hydrogen peroxide, the removal of bromine from $\beta\beta$-dibromo propionic acid and enables certain oxidations to take place in air at ordinary temperatures.

Warburg* studied the oxidation of oxalic- and of some α-amino-acids in dilute solution, in the presence of charcoal. He found that when he used blood charcoal, which contains iron, the reaction was inhibited by hydrocyanic acid in very low concentrations; if, however, he employed sugar charcoal, the rate of oxidation was not appreciably affected by small quantities of HCN. Sugar charcoal is itself slowly oxidised when it is suspended in water in contact with air. Warburg suggested that the catalytic action of blood charcoal was due to the iron, while the oxidation of acids by sugar charcoal was a coupled reaction with the auto-oxidation of the latter.

A more detailed investigation of oxidation at the surface of charcoals has now been made. The rate of oxidation was measured by a Barcroft differential manometer using 100 mgms. of charcoal in 3 c.c. of solution. In the first series of experiments Kahlbaum's sugar charcoal was used. It was previously activated by heating in silica vessels.

The rate of auto-oxidation was examined by shaking a suspension of the charcoal in conductivity water in a Barcroft adsorption vessel, in which the CO_2 evolved is adsorbed by a 40 per cent. solution of potash, so that the pressure change is proportional to the oxygen uptake. At 40°, the rate of oxygen uptake k is 0·13 mm.^{3} oxygen (N.T.P.) per milligram of charcoal per hour; at 50° the rate is 0·23. The temperature coefficient therefore is 1·80.

The reaction is of zero order, and independent of the pressure of the oxygen over the range examined. It therefore appears that the auto-oxidation takes place on very active parts of the charcoal surface, which are always saturated with oxygen; and the rate of reaction measured is equal to the rate of evaporation of carbon dioxide from them. Further, the combustion must proceed along a chain of carbon atoms, for, if it spread out on the surface in an expanding ring, the rate would continually increase, but this does not occur.

An attempt was made to measure the area of the auto-oxidisable surface by fractional poisoning with potassium ferricyanide, and with amyl alcohol.

* Warburg, *Pflüger's Arch.* vol. 55, p. 547 (1914).

Digitized by Google

The minimum concentration of poison required to arrest completely the auto-oxidation of the charcoal was determined, and the number of molecules adsorbed per milligram of charcoal, obtained from the adsorption isotherm; assuming that each molecule is adsorbed on one carbon atom, this gives the number of auto-oxidisable atoms per milligram. The total number of atoms on the surface per milligram was calculated from the adsorption maximum of amyl alcohol. The results showed that not more than 0·38 per cent. of the charcoal surface is auto-oxidisable.

The oxidation of oxalic acid to carbon dioxide and water on the surface of sugar charcoal is not coupled with the auto-oxidation of the charcoal, for the latter can be completely inhibited by the addition of small quantities of amyl alcohol, without appreciably affecting the rate of oxidation of the acid.

Warburg* states that the oxidation velocity of oxalic acid varies inversely as its bulk concentration. By starting with very dilute solutions, however, it was found that with increasing concentration, the velocity rises rapidly to a maximum, and falls off again. The position of this maximum is not appreciably altered by a rise in temperature of 10°, the temperature coefficient being 2·0 for all concentrations observed.

The maximum velocity is identical in air and in oxygen, in the latter, however, it is displaced towards higher oxalic acid concentrations, showing that at the higher pressure the acid is displaced from the surface by the oxygen, and a stronger solution is required in order that the two constituents may again be present in the optimum proportions.

By fractional poisoning with amyl alcohol it was shown that not more than 40 per cent. of the surface is catalytically acitive.

With Merk's blood charcoal, the maximum oxidation velocity of oxalic acid is some fifteen times that obtained with sugar charcoal. The specific surface of the former, as determined by adsorption of methylene blue, is eight times that of the latter.

The rate of oxidation is retarded by the addition of amyl alcohol, and the degree of poisoning is proportional to the fraction of surface covered up to 80 per cent. poisoning. Complete poisoning could not be obtained, owing to the slight solubility of amyl alcohol. By extrapolation, the catalytically acitive surface was estimated as 60 per cent. of the total surface.

If potassium cyanide is used as a poison, and the degree of poisoning plotted against the fraction of surface covered, it is found the reaction velocity is reduced to one-half of its original value, when only 1·5 per cent. of the surface is covered with adsorbed cyanide. At this point there is a distinct break in the curve, further addition of cyanide producing relatively little increase in

* Warburg, *Pflüger's Arch.* vol. 55, p. 547 (1914).

Digitized by Google

the amount of poisoning. At high concentrations there is an increase in the oxygen-uptake, owing to the oxidation of the KCN itself.

This shows that there are two catalytically active areas on blood charcoal, one small, highly active surface, which is preferentially poisoned by the cyanide ion, and another larger, less active surface. If the former is poisoned by the addition of $M/1000$ KCN, and the maximum velocity on the remaining surface determined, the specific activity of the two areas can be calculated.

The results show that the very active surface A oxidises $1{\cdot}1 \times 10^{12}$ mols oxalic acid per cm.^{2} surface per sec., while the active surface B oxidises $2{\cdot}8 \times 10^{10}$ mols oxalic acid per cm.^{2} surface per sec. The active surface of sugar charcoal oxidises $3{\cdot}6 \times 10^{10}$ mols per sec. The ratio of these activities $C:B:A$ is as 1·3:1:39·4.

Since the poisons used are not absolutely selective, the areas measured represent the maximum active areas and therefore the activities calculated from these are only minimum values. Areas C and B are probably identical, and may be attributed to the carbon surface. Surface A, which is about forty times as active as the former, probably consists of an iron-carbon union.

The temperature coefficient of the oxidation of oxalic acid on blood charcoal is $k\,(50°)/k\,(40°) = 1{\cdot}80$; this represents the resultant of the action on the two surfaces.

By finding the temperature coefficient, when the iron-containing area is poisoned with potassium cyanide, and correcting for the slight decrease in adsorption of the acid at the higher temperature, the temperature coefficient for each surface can be obtained. This gives for

surface A, $k\,(50°)/k\,(40°) = 1{\cdot}67$,
surface B, $k\,(50°)/k\,(40°) = 1{\cdot}93$,
surface C, $k\,(50°)/k\,(40°) = 2{\cdot}00$.

From these values the heat of activation Q is found to be for

surface A, 10,240 calories
surface B, 12,940 ,,
surface C, 13,660 ,,

Warburg* claims that only nitrogenous charcoals can be activated by iron. To test this point, some nitrogen-free charcoals were prepared containing small quantities of iron salts, together with control charcoals made without the addition of iron. A mixture of sugar and ferric chloride in a little water was heated in an atmosphere of carbon dioxide, until carbonisation was complete.

* Warburg, *Biochem. Zeits.* vol. 145, p. 461 (1924).

Digitized by Google

The resulting charcoal was baked out in a stream of hydrogen for six hours in order to remove the CO_2, and reduce the iron. The iron-containing charcoals were more active than the control specimens, and resembled blood charcoal, in that their catalytic activity was diminished by the addition of small quantities of potassium cyanide, and the specific activity of the two surfaces was of the same order in both charcoals.

Blood charcoal, surface A $= 3{\cdot}0 \times 10^{-5}$ c.c. O_2 per sq. cm.
Nitrogen-free charcoal, surface A' $= 4{\cdot}3 \times 10^{-5}$,, ,, ,,
Blood charcoal, surface B $= 7{\cdot}6 \times 10^{-7}$,, ,, ,,
Nitrogen-free charcoal, surface B' $= 9{\cdot}0 \times 10^{-7}$,, ,, ,,

Paneth and Radu's* data show that nitrogenous charcoals have a greater specific surface than those free from nitrogen; nitrogenous compounds appear to exert a peptising influence on the charcoal. This has now been shown by preparing a series of charcoals from mixtures of urea and sugar.

The specific surface of the charcoals increased with the urea content of the mixture up to about 5 per cent. urea, and the yield of active charcoal was greater than that obtained from sugar alone, for the nitrogenous charcoals are less easily oxidised, and thus there is less loss by combustion in the heating processes.

Percentage urea	Specific surface square metres per gram	Yield per 100 grams of sugar
0	22	8·0
0·5	42	9·8
5·0	84	12·2
10·0	83	15·0

Therefore, although iron can activate charcoal in the absence of nitrogen, yet nitrogenous charcoals are, in general, more efficient catalysts than pure sugar charcoals, owing to their greater specific surface; also, the ammonia produced during carbonisation of the former aids in reducing the iron; whereas the iron in sugar charcoal is inactive unless it is reduced in hydrogen. Further, some experiments now in progress indicate that an iron-carbon-nitrogen complex is a more active catalyst than the iron-carbon union.

In conclusion, I wish to thank Mr Rideal for suggesting this work, and for his interest and advice throughout its course.

* Paneth and Radu, *Ber.* vol. 57, p. 1221 (1924).

Digitized by Google

PROCEEDINGS
OF THE
Cambridge Philosophical Society.

The Problem of Relativity in reference to several bodies. By Mr R. HARGREAVES, St John's College.

[*Received* 23 January, *read* 1 March, 1926.]

§ 1. If the kinetic potential for the relative motion of two masses is written with an added constant as

$$2L = \frac{m_1 m_2}{m_1 + m_2}\left\{\dot{x}^2 + \dot{y}^2 + \dot{z}^2 - V^2 + \frac{2(m_1 + m_2)}{r}\right\},$$

a close connexion with the relativity quadratic appears. The latter is in fact

$$dx^2 + dy^2 + dz^2 + (\gamma^{-1} - 1)\,dr^2 - (V^2 - 2m/r)\,dt^2,$$

where $$\gamma = 1 - 2m/V^2 r;$$

a modification of the primary form

$$dx^2 + dy^2 + dz^2 - V^2 dt^2,$$

which shows an unaltered determinant. The condition in respect to the determinant, suggested, I believe, by Schwarzschild, is one which to me appears to give the most significant form to the results. From the dynamical standpoint we may regard it as imposing a counterpoise in the inertia coefficients to the modification introduced by the potential; or from a geometrical point of view we may regard it as minimizing the departure from the normal use of coordinates. An illuminating example of the loss of meaning that accompanies transformation in which this condition is disregarded is furnished by the isotropic form which is sometimes given to Einstein's quadratic.

It is proposed to examine whether it is practicable to apply a procedure suggested by the above considerations, when more than two bodies are concerned. The enquiry is prompted by a conviction that any attempt to modify $dx^2 + dy^2 + dz^2 - V^2 dt^2$ by use of a form based on accelerations to two or more *fixed* centres is inadmissible, on the ground that it fails to acknowledge the essential dynamical relativity of the bodies; or in other words the method is open to the objection that it does not deal with all bodies on equal terms, another way of stating the breach of the principle of relativity involved.

Digitized by Google

It appears to me necessary to use a space quadratic* which shall include all relative coordinates, and one time coordinate. For this purpose one of the masses m is chosen arbitrarily as a centre from which rectangular coordinates are measured in fixed directions. Thus another mass m_p has coordinates x_p, y_p, z_p relative to m, and M may be used for the sum of masses $m + m_1 + m_2 + \ldots$. The primary form for several bodies is then taken to be

$$\left.\begin{aligned} & mm_1M^{-1}(dx_1^2 + dy_1^2 + dz_1^2 - V^2dt^2) \\ & \qquad + mm_2M^{-1}(dx_2^2 + dy_2^2 + dz_2^2 - V^2dt^2) + \ldots \\ & + m_1m_2M^{-1}\{(dx_2 - dx_1)^2 + (dy_2 - dy_1)^2 \\ & \qquad + (dz_2 - dz_1)^2 - V^2dt^2\} + \ldots \end{aligned}\right\} \quad (1).$$

As a first step terms

$$2dt^2(mm_1r_1^{-1} + mm_2r_2^{-1} + \ldots + m_1m_2r_{12}^{-1} + \ldots)$$

are added, which put us in touch with the kinetic potential for the relative motion of several bodies. It is then assumed that terms added to inertia take the shapes $\mu_1 dr_1^2/r_1$ or $\mu_{12}dr_{12}^2/r_{12}$ in the respective brackets of (1), and we seek to ascertain whether the values of the μ's admit of an adjustment which will make the inertial terms act as a counterpoise to potential terms, in the sense that the determinant of the modified quadratic agrees with that of the original. In using $\mu dr^2/r$ we have in mind the approximate value of $\gamma^{-1} - 1$ above, viz. $2m/rV^2$; our main concern is with a first approximation, though the exact problem will be briefly considered.

§ 2. The method will be applied to the case of four bodies m, m_1, m_2, m_3; the argument is quite general, but can be conducted with a less cumbrous notation than that required for n bodies. The coordinates $x_1, y_1, z_1 \ldots z_3$ are counted as having places 1, 2, 3 ... 9 in the scheme of space variables, and $a_{11} \ldots a_{99}$ are the elements of the space-determinant of the *unmodified* form (1). The property $a_{rs} = a_{sr}$ is general, and the special values of the a's for four bodies are:

$$\left.\begin{aligned} a_{11} = a_{22} = a_{33} &= m_1(M - m_1)/M = a \text{ say} \\ a_{44} = a_{55} = a_{66} &= m_2(M - m_2)/M = b \text{ say} \\ a_{77} = a_{88} = a_{99} &= m_3(M - m_3)/M = c \text{ say} \\ a_{14} = a_{25} = a_{36} &= -m_1m_2/M = h \text{ say} \\ a_{47} = a_{58} = a_{69} &= -m_2m_3/M = f \text{ say} \\ a_{71} = a_{82} = a_{93} &= -m_3m_1/M = g \text{ say} \end{aligned}\right\} \quad \ldots\ldots(2).$$

* The need for recognition of change in all space variables is emphasized by M. Le Roux, as I see in a quotation from *Comptes Rendus* given by Sir Joseph Larmor, who also presses the same point, *Proc. Camb. Phil. Soc.*, vol. XXI, pp. 415 *sqq.*

Digitized by Google

With other suffixes the a's have a zero value, as a_{12} or a_{21}, since $dx\,dy$ does not occur. The determinant D of the space-quadratic is then

$$\begin{vmatrix} a_{11} & 0 & 0 & a_{14} & 0 & 0 & a_{17} & 0 & 0 \\ 0 & a_{22} & 0 & 0 & a_{25} & 0 & 0 & a_{28} & 0 \\ 0 & 0 & a_{33} & 0 & 0 & a_{36} & 0 & 0 & a_{39} \\ a_{41} & 0 & 0 & a_{44} & 0 & 0 & a_{47} & 0 & 0 \\ 0 & a_{52} & 0 & 0 & a_{55} & 0 & 0 & a_{58} & 0 \\ 0 & 0 & a_{63} & 0 & 0 & a_{66} & 0 & 0 & a_{69} \\ a_{71} & 0 & 0 & a_{74} & 0 & 0 & a_{77} & 0 & 0 \\ 0 & a_{82} & 0 & 0 & a_{85} & 0 & 0 & a_{88} & 0 \\ 0 & 0 & a_{93} & 0 & 0 & a_{96} & 0 & 0 & a_{99} \end{vmatrix}$$

which is equal to the product

$$\begin{vmatrix} a_{11} & a_{14} & a_{17} \\ a_{41} & a_{44} & a_{47} \\ a_{71} & a_{74} & a_{77} \end{vmatrix} \times \begin{vmatrix} a_{22} & a_{25} & a_{28} \\ a_{52} & a_{55} & a_{58} \\ a_{82} & a_{85} & a_{88} \end{vmatrix} \times \begin{vmatrix} a_{33} & a_{36} & a_{39} \\ a_{63} & a_{66} & a_{69} \\ a_{93} & a_{96} & a_{99} \end{vmatrix} \quad (3).$$

The equality depends on the places of zero values in D, not on the special values shown in (2). The first minors of D are given by

$$\left.\begin{aligned} D_{11} &= \begin{vmatrix} a_{44} & a_{47} \\ a_{74} & a_{77} \end{vmatrix} \times \begin{vmatrix} a_{22} & a_{25} & a_{28} \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{vmatrix} \times \begin{vmatrix} a_{33} & a_{36} & a_{39} \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{vmatrix} \\ D_{41} &= -\begin{vmatrix} a_{14} & a_{17} \\ a_{74} & a_{77} \end{vmatrix} \times \begin{vmatrix} a_{22} & a_{25} & a_{28} \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{vmatrix} \times \begin{vmatrix} a_{33} & a_{36} & a_{39} \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{vmatrix} \\ D_{71} &= \begin{vmatrix} a_{14} & a_{17} \\ a_{44} & a_{47} \end{vmatrix} \times \begin{vmatrix} a_{22} & a_{25} & a_{28} \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{vmatrix} \times \begin{vmatrix} a_{33} & a_{36} & a_{39} \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{vmatrix} \end{aligned}\right\} \quad (4).$$

For the values of D_{22}, D_{52}, D_{82} we take minors of the middle third-order determinant in (3) and associate them with the product of the other two. All first minors of D which correspond to zero places in D vanish, as $D_{21} = 0$.

§ 3. When account is taken of such terms as $\mu_1 dr_1^2/r_1$ in the modified form, the coefficients are more complicated, and none have zero value nor do any first minors vanish. But since a minor such as D_{21} is zero when μ's are absent, it follows that a term $a_{21}D_{21}$ is of the second order of small quantities. Therefore in the approximate evaluation of the determinant we may count places which had zero values as still zero, while taking account of modifications in the a's and in first minors for terms which had value in the original scheme.

Digitized by Google

Thus to trace the effect of μ_1 we note that the addition to the quadratic is

$$mm_1 M^{-1}\mu_1 dr_1^2/r_1$$

or $$mm_1 M^{-1}\mu_1 (x_1 dx_1 + y_1 dy_1 + z_1 dz_1)^2/r_1^3,$$

and so with dashed letters for modified values

$$\left.\begin{aligned} a_{11}' &= a_{11} + mm_1 M^{-1}\mu_1 x_1^2/r_1^3 \\ a_{22}' &= a_{22} + mm_1 M^{-1}\mu_1 y_1^2/r_1^3 \\ a_{33}' &= a_{33} + mm_1 M^{-1}\mu_1 z_1^2/r_1^3 \end{aligned}\right\} \quad \text{.....(5),}$$

and it is not necessary to write a_{21}' .... The first of the product determinants in (3) is replaced by

$$\begin{vmatrix} a + mm_1 M^{-1}\mu_1 x_1^2/r_1^3 & h & g \\ h & b & f \\ g & f & c \end{vmatrix}$$

which is $$\Delta + mm_1 M^{-1}\mu_1 x_1^2 r_1^{-3} (bc - f^2) \quad \text{.....(6),}$$

where Δ is the original form of each of the third order determinants. With the values in (2)

$$\Delta = mm_1 m_2 m_3/M, \qquad bc - f^2 = m_2 m_3 (m + m_1)/M,$$

and therefore (6) is $\Delta\left(1 + \dfrac{m + m_1}{M} \cdot \dfrac{\mu_1 x_1^2}{r_1^3}\right)$.

The second and third of the determinants in (2) differ only in having y_1^2 and z_1^2 in lieu of x_1^2, and accordingly the value of D as modified by μ_1 only is

$$\Delta^3\left(1 + \frac{m + m_1}{M} \cdot \frac{\mu_1}{r_1}\right).$$

There is complete symmetry in the quadratic form, and therefore, so long as we attend only to first order corrections, the whole modified determinant is

$$\Delta^3\left(1 + \frac{m + m_1}{M}\frac{\mu_1}{r_1} + \ldots + \frac{m_1 + m_2}{M}\frac{\mu_{12}}{r_{12}} + \ldots\right) \quad \text{.....(7).}$$

A little more algebra is needed to reach the final form in the case of μ_{12}, but in view of the manifest symmetry it seemed sufficient to take the simplest case.

If now we write P for the sum of products of masses two and two, $mm_1 + \ldots + m_1 m_2 + \ldots$, the increment in the determinant of *tenth* order where time is included, due to the change in the spatial section, is

$$-\frac{PV^2}{M} \times \Delta^3\left(\frac{m + m_1}{M}\frac{\mu_1}{r_1} + \ldots\right),$$

while the increment due to the potential terms which modify the coefficient of dt^2 is

$$2\Delta^3\left(\frac{mm_1}{r_1} + \ldots\right).$$

Digitized by Google

When these are made to balance we have

$$\mu_1 V^2 = 2mm_1 M^2/(m + m_1)\, P, \quad \mu_{12} V^2 = 2m_1 m_2 M^2/(m_1 + m_2)\, P, \ldots \quad \ldots\ldots(8).$$

Thus with a choice of values for the μ's dependent solely on the masses concerned, the modified determinant of the tenth order retains the value of the original. It will be observed that in the formula for μ_{12} other masses than m_1 and m_2 enter only through the values of M and P; if only m and m_1 exist then μ_1 becomes $2(m + m_1)/V^2 r$, the primary case from which we set out. Also it may be well to state, in respect to the problem for any number of bodies, that the determinant of the unmodified space quadratic is always the cube of a constituent Δ, the order of which is one less than the number of bodies. For three bodies it has a form

$$\begin{vmatrix} a_{11} & a_{14} \\ a_{41} & a_{44} \end{vmatrix}$$

equal in value to $\begin{vmatrix} a_{22} & a_{25} \\ a_{52} & a_{55} \end{vmatrix}$ or $\begin{vmatrix} a_{33} & a_{36} \\ a_{63} & a_{66} \end{vmatrix}$,

while for five bodies Δ has a first row

$$a_{11} \quad a_{14} \quad a_{17} \quad a_{1,10}.$$

§ 4. The question is now what treatment is to be applied to the modified quadratic we have reached, as for example to that of the seventh order required for the relative motion of three bodies. As I understand the matter, the primary treatment is to take

$$dx_1/ds, \ldots d(Vt)/ds$$

as velocities, and apply the usual variation method to a kinetic potential

$$L = \frac{mm_1}{2M}\left\{\left(\frac{dx_1}{ds}\right)^2 + \ldots\right\} + \ldots$$

written with all the modifying terms. This is a purely kinetic system with seven velocities, and t is an ignorable coordinate.

I have given elsewhere* a method which combines the ignoration with a transfer to the use of the ignored coordinate as independent variable. The application of this method yields a new kinetic potential

$$L = \frac{mm_1}{2M\gamma}(\dot{x}_1^2 + \dot{y}_1^2 + \dot{z}_1^2 + \mu_1 \dot{r}_1^2/r_1) + \ldots - \frac{PV^2}{2M} + \frac{mm_1}{r_1} + \frac{mm_2}{r_2} + \frac{m_1 m_2}{r_{12}} \quad \ldots\ldots(9),$$

* *Proc. Camb. Phil. Soc.*, vol. XXII, Part 2, § 17. There is a divisor h^2 affecting the gravitation constant, where h is the constant value of $\gamma d(Vt)/ds$; in the above this is taken to be 1, or to be absorbed in the gravitation constant.

Digitized by Google

$\dot{x}$ being dx/dt and

$$\gamma = 1 - \frac{2M}{PV^2}\left(\frac{mm_1}{r_1} + \frac{mm_2}{r_2} + \frac{m_1 m_2}{r_{12}}\right).$$

Or we may say that the usual formula

$$\frac{mm_1}{2M}(\dot{x}_1^2 + \dot{y}_1^2 + \dot{z}_1^2) + \ldots + \frac{mm_1}{r_1} + \frac{mm_2}{r_2} + \frac{m_1 m_2}{r_{12}}$$

is to be supplemented by first-order modifications

$$\frac{1}{2M}(mm_1\mu_1\dot{r}_1^2/r_1 + mm_2\mu_2\dot{r}_2^2/r_2 + m_1 m_2 \dot{r}_{12}^2/r_{12})$$
$$+ \frac{1}{PV^2}\{mm_1(\dot{x}_1^2 + \dot{y}_1^2 + \dot{z}_1^2) + \ldots\}\left(\frac{mm_1}{r_1} + \frac{mm_2}{r_2} + \frac{m_1 m_2}{r_{12}}\right) \quad (10).$$

Thus the method, as here carried through, leads to a definite modification of the normal kinetic potential belonging to the problem of n bodies; and, as noted at the end of § 3, the supplementary terms are, in the case of two bodies only, reduced to the corrections given by Einstein.

A brief summary of the argument may be added. From the dynamical point of view it is a striking feature of Einstein's quadratic form that it provides by ignoration of t the potential terms of dynamics. It is assumed that this is also the case with several bodies, so that the content of a modified coefficient of dt^2 is dictated by our knowledge of the potential. By analogy with the case of two bodies compensatory terms varying as dr^2 are introduced, and if the type is admitted the coefficients follow from the principle that the determinant is unchanged in value.

§ 5. Curiosity led me to examine, in the case of three bodies, the exact value of the spatial determinant. For that purpose the quadratic may be written as

$$A(dx_1^2 + dy_1^2 + dz_1^2) + B(dx_2^2 + dy_2^2 + dz_2^2)$$
$$+ C\{(dx_2 - dx_1)^2 + \ldots\} + a\,dr_1^2 + b\,dr_2^2 + c\,dr_{12}^2 \quad \ldots(11).$$

With $dr_1 = l_1 dx_1 + m_1 dy_1 + n_1 dz_1$, $\quad dr_{12} = l_3(dx_2 - dx_1) + \ldots$

we have three systems of direction-cosines connected by

$$l_3 r_{12} = x_2 - x_1 = l_2 r_2 - l_1 r_1, \ldots.$$

The determinant is then

$$\begin{vmatrix} A + C + al_1^2 + cl_3^2, & al_1m_1 + cl_3m_3, & al_1n_1 + cl_3n_3, & -C - cl_3^2, & -cl_3m_3, & -cl_3n_3 \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ -C - cl_3^2, & -cl_3m_3, & -cl_3n_3, & B + C + bl_2^2 + cl_3^2, & bl_2m_2 + cl_3m_3, & bl_2n_2 + cl_3n_3 \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \end{vmatrix}$$

$\ldots\ldots(12).$

Digitized by Google

The part independent of a, b, c is $(AB + BC + CA)^3$ or Δ^3. Each of the letters a, b, c may occur in the first but in no higher power; the value of the determinant is

$$\left.\begin{aligned}&\Delta^3 + \Delta^2\{a(B+C) + b(C+A) + c(A+B)\} + \Delta^2(ab+bc+ca)\\&\quad + 4\Delta S^2(abC^2r_1^{-2}r_2^{-2} + bcA^2r_2^{-2}r_{12}^{-2} + caB^2r_1^{-2}r_{12}^{-2})\\&\quad + 4\Delta S^2abc(Cr_1^{-2}r_2^{-2} + Ar_2^{-2}r_{12}^{-2} + Br_1^{-2}r_{12}^{-2}).\end{aligned}\right\} \quad (13).$$

The direction-cosines occur in combinations which may be expressed in terms of the mutual distances. There is, for example, a term

$$ab\Delta\{(A+C)(B+C) - C^2(\Sigma l_1 l_2)^2\}$$

which is

$$ab\Delta(\Delta + 4C^2S^2r_1^{-2}r_2^{-2}),$$

where S is the area of the triangle formed by the three point-masses.

We may alter the notation to bring it into closer relation with previous work by setting $a = A\lambda_1$, $b = B\lambda_2$, $c = C\lambda_{12}$, so that λ_1 takes the place of μ_1/r_1. Since A, B, C, Δ have the values mm_1M^{-1}, mm_2M^{-1}, $m_1m_2M^{-1}$, $mm_1m_2M^{-1}$ the value of the determinant (12) is

$$\left.\begin{aligned}&\Delta^3[1 + M^{-1}\{(m+m_1)\lambda_1 + (m+m_2)\lambda_2 + (m_1+m_2)\lambda_{12}\}\\&\quad + M^{-1}(m\lambda_1\lambda_2 + m_1\lambda_1\lambda_{12} + m_2\lambda_2\lambda_{12})\\&\quad + 4S^2M^{-2}(m_1m_2\lambda_1\lambda_2r_1^{-2}r_2^{-2} + mm_1\lambda_2\lambda_{12}r_2^{-2}r_{12}^{-2}\\&\quad + mm_2\lambda_1\lambda_{12}r_1^{-2}r_{12}^{-2}) + 4S^2M^{-2}\lambda_1\lambda_2\lambda_3(m_1m_2r_1^{-2}r_2^{-2}\\&\quad + mm_1r_2^{-2}r_{12}^{-2} + mm_2r_1^{-2}r_{12}^{-2})]\end{aligned}\right\} \quad (14).$$

The property of the determinant that it depends solely on the mutual distances is interesting, and may probably be true for any number of bodies. I have no speculation to suggest in reference to terms beyond the first order; in the inertia section they can hardly be within the limits of observation, in the potential section they might be.

Digitized by Google

Relativity in Connexion with Axial Rotation. By Mr R. HARGREAVES, St John's College.

[*Received* 25 January, *read* 1 March, 1926.
Received in revised form 4 May 1926.]

§ 1. My primary aim here is to give some account of relativity in connexion with axial rotation, and in particular to deal with the disk problem. The justification of the use of a special quadratic form for this purpose suggests the propriety of some introductory remarks of a more general character.

Einstein's formula $\gamma^{-1}dr^2 + r^2d\theta^2 + r^2\sin^2\theta d\phi^2 - \gamma dt^2$ leads to values of the covariant derivatives B (in his notation) given by

$$B_{11}/g_{11} = B_{44}/g_{44} = \frac{1}{2r^2}\frac{d}{dr}\left(r^2\frac{d\gamma}{dr}\right) \qquad \text{.....(1).}$$

It is clear then that the use of a zero or a constant value on the left hand leads to the special forms of Laplace's or Poisson's equation which belong to the case where r is the only variable. Thus with $\gamma = 1 - 2m/r$, an external value, we may associate an internal value $\gamma = 1 - 4\pi\rho a^2 + 4\pi\rho r^2/3$, continuous at $r = a$ with the first and also having $\frac{d\gamma}{dr}$ continuous, when $m = 4\pi\rho a^3/3$.

For the other B's we have

$$B_{22} = \frac{d}{dr}(r\gamma) - 1, \text{ and } B_{33} = B_{22}\sin^2\theta \qquad \text{.....(2);}$$

and the substitution of the above internal value for γ does not make B_{22}/g_{22} constant. If we choose to drop the condition of continuity for γ, we may retain continuity for $\frac{d\gamma}{dr}$ by writing $\gamma = 1 - 2m/r$ and $\gamma = 1 + 4\pi\rho r^2/3$; that is, continuity in respect to acceleration remains, but an element of connexion with gravitational potential is lost. Thus a question arises whether the quadratic forms have reference to a spherical attracting *body*, with adaptation to internal and external positions, or simply to two different laws of central acceleration.

If it were supposed for example that there exist quadratic forms correlated with the attraction of a series of spheroids, the analytical test would be more serious than for the sphere, but the limit of prolate spheroids, viz. the infinite cylinder, is a case where the analysis is as simple as for the sphere, and there is no likelihood

Digitized by Google

that a critical position is involved in the mere prolongation of an axis of the spheroid.

Now with a cylindrical form $\gamma^{-1}d\varpi^2 + \varpi^2 d\phi^2 + dz^2 - \gamma dt^2$, where γ depends only on ϖ, we have

$$B_{11}/g_{11} = B_{44}/g_{44} = \frac{1}{2\varpi}\frac{d}{d\varpi}\left(\varpi\frac{d\gamma}{d\varpi}\right), \quad B_{22} = \varpi\frac{d\gamma}{d\varpi}, \quad B_{33} = 0 \quad (3).$$

With respect to B_{11} and B_{44} the position in respect to continuity is the same as for the sphere with values $\gamma = 1 + 4\pi\rho a^2 \log \varpi$ and $\gamma = 1 + 2\pi\rho\varpi^2 + 4\pi a^2 \log a - 2\pi\rho a^2$. But the first of these gives $B_{22} = 4\pi\rho a^2$ instead of 0, and the second gives $B_{33} = 0$ instead of $4\pi\rho$.

The first discrepancy cannot be met, and I take it there is no gravitational form for the cylinder. The second discrepancy can be met by taking $g_{33} = 0$, that is by treating the problem as strictly two-dimensional; but with the failure of the external case this naturally goes with it. The form may be interpreted with a different constant as applying to axial rotation (*v. infra*). It may be observed that the values of B's are not altered by writing $g_{33} = 0$ for $g_{33} = 1$.

If we then abandon the reference of one type of solution to space occupied by matter and interpret all as belonging to different laws of acceleration in free space, the position is as follows. There is (i) in three dimensions a quadratic form corresponding to an imposed acceleration m/r^2, in which $\gamma = 1 - 2m/r$. Again in three dimensions there is (ii) a form corresponding to an imposed acceleration μr, where $\gamma = 1 + \mu r^2$. In two dimensions there is (iii) a form corresponding to an inertial acceleration $\omega^2\varpi$, in which $\gamma = 1 - \omega^2\varpi^2$; and here $B_{pp} = -2\omega^2 g_{pp}$. Thus the condition $B_{pp} = 0$ attaches to acceleration diminishing with increasing distance, the condition $B_{pp} = g_{pp}$ to the reverse case.

§ 2. We proceed to the consideration of this last form which, with V written explicitly, is

$$V^2 d\varpi^2/(V^2 - \omega^2\varpi^2) + \varpi^2 d\phi^2 - (V^2 - \omega^2\varpi^2)\, dt^2.$$

By way of introduction we find a radius r defined as a geodetic length from the origin when $d\phi$ and dt are taken to be zero. It is given by

$$r = \int ds = \int_0^{\varpi} \frac{V\, d\varpi}{\sqrt{V^2 - \omega^2\varpi^2}} = \frac{V}{\omega}\sin^{-1}\frac{\omega\varpi}{V}, \text{ or } \frac{\omega\varpi}{V} = \sin\frac{r\omega}{V} \quad (4).$$

The area in which ϕ makes a complete circuit is

$$A = \iint \frac{V\varpi\, d\varpi\, d\phi}{\sqrt{V^2 - \omega^2\varpi^2}} = \frac{2\pi V}{\omega^2}(V - \sqrt{V^2 - \omega^2\varpi^2}) = \frac{2\pi V^2}{\omega^2}\left(1 - \cos\frac{r\omega}{V}\right)$$

.....(5).

Digitized by Google

The circumference being $C = 2\pi\varpi$, we have a relation

$$C^2/A = 2\pi\,(V + \sqrt{V^2 - \omega^2\varpi^2})/V = 4\pi\cos^2 r\omega/2V \quad ...(6).$$

The extreme value possible for ϖ is V/ω; with this value

$$r = \pi V/2\omega, \text{ and } C^2/A = 2\pi \quad .....(7)$$

in lieu of the normal value 4π.

Comparison with measures in a non-rotating case involves an arbitrary element; perhaps the most significant is that in which areas are made equal.

Thus if we write

$$\left.\begin{aligned} \pi a^2 = 2\pi V\,(V - \sqrt{V^2 - \omega^2\varpi^2})/\omega^2 = 2\pi V^2\,(1 - \cos r\omega/V)/\omega^2 \\ \text{we find } \varpi^2 = a^2\,(1 - \omega^2 a^2/4V^2), \quad a\omega/2V = \sin r\omega/2V, \\ C = 2\pi a\sqrt{1 - \omega^2 a^2/4V^2} \end{aligned}\right\} \quad (8).$$

An assumed equality of areas makes $C < 2\pi a$, while making $r > a$.

The above formulae suggest a general treatment of geodetic lengths and directions when $dt = 0$. If we write $R = V/\omega$, and $\varpi = R\sin\theta$, we have

$$ds^2 = V^2 d\varpi^2/(V^2 - \omega^2\varpi^2) + \varpi^2 d\phi^2 = R^2\,(d\theta^2 + \sin^2\theta\, d\phi^2) \quad (9).$$

Thus a complete scheme of geodesy for the rotating disk is furnished by reference to the spherical trigonometry of a hemisphere of radius V/ω. For example, if lines OP_1, OP_2 of geodetic lengths $r_1 r_2$ are drawn from the centre at an angle ϕ, the length of the geodetic line from P_1 to P_2 is given by

$$\cos\frac{\omega s}{V} = \cos\frac{\omega r_1}{V}\cos\frac{\omega r_2}{V} + \sin\frac{\omega r_1}{V}\sin\frac{\omega r_2}{V}\cos\phi \ldots\ldots(10),$$

and its course is given by the great circle of the sphere which joins the two representative points.

§ 3. The plane motion of a particle in the scheme of (3) is got by finding a geodetic path when ϖ, ϕ and t all vary. With $\varpi^2\dot{\phi} = k\gamma$, $\gamma = 1 - \omega^2\varpi^2/V^2$, and $\varpi u = 1$ the orbit is given by an equation

$$k^2\left\{\left(\frac{du}{d\phi}\right)^2 + \gamma u^2\right\} - \omega^2\varpi^2 = \text{constant},$$

which only differs from a corresponding result for Einstein's central acceleration problem in having $\omega^2\varpi^2$ for $2m/r$, so that it is not necessary to give details. Thus we have

$$\left(\frac{du}{d\phi}\right)^2 + u^2\left(1 - \frac{\omega^2\varpi^2}{V^2}\right) - \frac{\omega^2\varpi^2}{k^2} = \text{constant},$$

or

$$\left(\frac{du}{d\phi}\right)^2 + u^2 - \frac{\omega^2}{k^2 u^2} = \text{constant}.$$

Digitized by Google

We may therefore write

$$u^2\left(\frac{du}{d\phi}\right)^2 = (\alpha^2 - u^2)(\beta^2 + u^2), \text{ where } \alpha^2\beta^2 = \frac{\omega^2}{k^2} \quad ...(11).$$

A solution is $u^2 = \alpha^2\cos^2\phi - \beta^2\sin^2\phi$, that is with

$$x = \varpi\cos\phi, \quad y = \varpi\sin\phi, \quad 1 = \alpha^2x^2 - \beta^2y^2 .....(12).$$

The course of the curve is not completed, but only followed to limits imposed by $V^2 = \omega^2\varpi^2$ or $\omega^2/V^2 = \alpha^2\cos^2\phi_0 - \beta^2\sin^2\phi_0$, that is up to $\tan^2\phi_0 = (V^2\alpha^2 - \omega^2)/(V^2\beta^2 + \omega^2) < \alpha^2/\beta^2$. When ω/α and ω/β are small compared with V a considerable approach to asymptotic directions can be made. The addition of a phase-constant to ϕ gives the general solution

$$\alpha^2\cos^2(\phi + \epsilon) - \beta^2\sin^2(\phi + \epsilon) = u^2 = 1/\varpi^2.$$

In the treatment of a ray of light in two dimensions the problem is defined by

$$\dot{\varpi}^2/\gamma^2 + \varpi^2\dot{\phi}^2/\gamma = V^2, \text{ and } \varpi^2\dot{\phi} = k\gamma \quad .....(13),$$

leading to

$$\left(\frac{du}{d\phi}\right)^2 + u^2\left(1 - \frac{\omega^2\varpi^2}{V^2}\right) = \frac{V^2}{k^2}, \text{ or } \left(\frac{du}{d\phi}\right)^2 + u^2 = \frac{V^2}{k^2} + \frac{\omega^2}{V^2} \quad (14).$$

The solution is a linear equation in x and y, defined as $\varpi\cos\phi$ and $\varpi\sin\phi$; this does not represent a geodesic for rest. Also the whole length of the line is not traversed; for example, if

$$a = \varpi\cos\phi = x$$

is the line, values of y extend to limits given by $V^2 = \omega^2(a^2 + y^2)$.

§ 4. In Eddington's *Relativity* there is a short account of the disk problem based on a transformation, which in simpler form is the substitution of $\phi + \omega t$ for θ, viz.

$$d\varpi^2 + \varpi^2d\theta^2 - V^2dt^2 = d\varpi^2 + \varpi^2d\phi^2 + 2\omega\varpi^2d\phi\,dt - (V^2 - \omega^2\varpi^2)\,dt^2.$$

The transformation seems to be used only to give the coefficient of dt^2, and a new value is taken for ϖ whereas the transformation leaves ϖ unaltered. Further, I have failed to follow the argument in reference to "proper" volume. In a quadratic where no terms containing products of dt and space-differentials occur, the whole determinant is the product of g_{44} and a space determinant; a certain separateness is thus given to volume and time elements, which are interlocked where product terms exist. In the above quadratic such a term exists, the space determinant is ϖ^2, the coefficient of dt^2 is $-(V^2 - \omega^2\varpi^2)$, and the product of the two is not the value of the whole determinant, which is $-V^2\varpi^2$. These difficulties led me to the conviction that a satisfactory account could only be reached by the application of a true acceleration form.

Digitized by Google

The existence of the discrepancies noted in § 1 was pointed out in an earlier paper*. The consideration of these in conjunction with the need for recognizing acceleration as a primary feature in dealing with axial rotation led to the formulation in § 1, as required to give proper significance to the part played by the B covariants.

* *Proc. Camb. Phil. Soc.* vol. XXII, Part 2, p. 158.

Digitized by Google

Sur la non-existence simultanée de deux fonctions de Mathieu. Par Ž. MARKOVIĆ. (Communicated by Mr E. L. INCE.)

[*Received* 1 April, *read* 3 May, 1926.]

1. Soit donnée l'équation de Mathieu

$$\frac{d^2y}{dx^2} + (a + 16q\cos 2x)\, y = 0, \tag{1·1}$$

a et q étant des paramètres réels.

Soient $s(x)$ et $c(x)$ deux solutions particulières de cette équation définies par les conditions initiales

$$s(0) = 0, \qquad c(0) = 1,$$
$$\dot{s}(0) = 1, \qquad \dot{c}(0) = 0.$$

$s(x)$ est une fonction impaire, $c(x)$ une fonction paire formant un système fondamental de l'équation (1·1) et satisfaisant à l'identité

$$c(x)\,\dot{s}(x) - s(x)\,\dot{c}(x) = 1. \tag{1·2}$$

Les fonctions $s(\pi + x)$, $s(\pi - x)$, $c(\pi + x)$, $c(\pi - x)$ étant aussi des solutions de l'équation (1·1), c'est à dire des combinaisons linéaires à coefficients constants des fonctions $s(x)$ et $c(x)$, on a, en déterminant convenablement ces coefficients, les relations suivantes:

$$\begin{aligned} s(\pi + x) &= c(x)\, s(\pi) + s(x)\, \dot{s}(\pi), \\ s(\pi - x) &= c(x)\, s(\pi) - s(x)\, \dot{s}(\pi), \\ c(\pi + x) &= c(x)\, c(\pi) + s(x)\, \dot{c}(\pi), \\ c(\pi - x) &= c(x)\, c(\pi) - s(x)\, \dot{c}(\pi). \end{aligned} \tag{1·3}$$

En faisant dans la dernière $x = \pi$, on a

$$c^2(\pi) = 1 + s(\pi)\, \dot{c}(\pi), \tag{1·4}$$

qui comparée à (1·2) pour $x = \pi$ donne la relation connue

$$\dot{s}(\pi) = c(\pi). \tag{1·5}$$

2. Pour obtenir la formule dont nous aurons besoin pour la démonstration du théorème de M. Ince sur l'impossibilité de coexistence de deux fonctions de Mathieu, multiplions l'équation

$$\ddot{s} + \phi(x)\, s = 0$$

par $\dot{c}$, et l'équation $\qquad \ddot{c} + \phi(x)\, c = 0$

par $\dot{s}$, où $\phi(x) = a + 16q\cos 2x$, et intégrons la somme des produits ainsi obtenus.

Digitized by Google

En tenant compte des conditions initiales de $s(x)$ et $c(x)$, on obtient après une intégration par parties la formule voulue:

(2·1)
$$\dot{s}(x)\,\dot{c}(x) + \phi(x)\,s(x)\,c(x) + 32q\int_0^x s(x)\,c(x)\sin 2x\,dx = 0.$$

Pour les valeurs de a et q qui correspondent à la fonction périodique $s(x)$ de période π, c'est à dire satisfaisant aux conditions
$$s(\pi) = 0, \qquad \dot{s}(\pi) = 1,$$
on a de (2·1)

(2·2) $$\dot{c}(\pi) + 32q\int_0^\pi s(x)\,c(x)\sin 2x\,dx = 0.$$

Pour l'existence simultanée de la fonction périodique $c(x)$ de période π, il faut que l'on ait en plus et pour les mêmes valeurs de a et q
$$c(\pi) = 1, \qquad \dot{c}(\pi) = 0,$$
ce qui d'après (1·5) et en vertu de (2·2) se réduit pour $q \neq 0$ à la seule condition:
$$\int_0^\pi s(x)\,c(x)\sin 2x\,dx = 0,$$
dont nous allons démontrer l'impossibilité.

Car si l'on pose
$$I_n = \int_0^\pi s(x)\,c(x)\sin 2nx\,dx,$$
en effectuant trois intégrations par parties successives sur les quantités I_n en supposition de $s(x)$ et $c(x)$ périodiques simultanément et en ayant égard au fait que la fonction $u(x) = s(x)\,c(x)$ satisfait à l'équation différentielle du troisième ordre
$$\dddot{u} + 4\phi(x)\,\dot{u} + 2\dot{\phi}(x)\,u = 0,$$
on trouve pour les quantités I_n la relation de récurrence
$$qI_{n+1} = \frac{n(n^2-a)}{4(2n+1)}I_n - q\frac{2n-1}{2n+1}I_{n-1}.$$

Si $I_1 = \int_0^\pi s(x)\,c(x)\sin 2x\,dx$ était nul, comme $I_0 = 0$, tous les I_n seraient nuls. Mais, dans notre hypothèse, la fonction impaire $u(x)$ est périodique de période π, les quantités I_n étant à un facteur constant près ses coefficients de Fourier; elle aurait donc tous ses coefficients de Fourier nuls, c'est qui est impossible.

3. Le cas des fonctions périodiques $s(x)$ et $c(x)$ de période 2π se ramène au précédent. En posant $c(\pi) = -1$, $s(\pi) = 0$ dans la

Digitized by Google

première des relations (1·3) et $c(\pi) = -1$, $\dot{c}(\pi) = 0$ dans la troisième, on a

$$s(\pi + x) = -s(x), \tag{3·1}$$
$$c(\pi + x) = -c(x),$$

c'est à dire

$$s(2\pi + x) = s(x),$$
$$c(2\pi + x) = c(x).$$

Les relations (3·1) sont donc valables pour les fonctions $s(x)$ et $c(x)$ périodiques de période 2π.

Mais dans l'hypothèse de la périodicité simultanée des fonctions $s(x)$ et $c(x)$ on a d'après (3·1)

$$s(\pi + x)\, c(\pi + x) = s(x)\, c(x);$$

la fonction impaire $u(x) = s(x)\, c(x)$ admet donc toujours la période π.

D'autre part, par un raisonnement analogue à celui du n°. 2, on conclut que pour l'existence simultanée des fonctions $s(x)$ et $c(x)$ de période 2π on devrait avoir pour $q \neq 0$

$$\int_0^{2\pi} s(x)\, c(x) \sin 2x\, dx = 0.$$

En remplaçant dans l'intervalle $(\pi, 2\pi)$ x par $\pi + x$, on trouve en vertu de (3·1)

$$\int_0^{2\pi} s(x)\, c(x) \sin 2x\, dx = 2 \int_0^{\pi} s(x)\, c(x) \sin 2x\, dx;$$

on devrait donc avoir

$$\int_0^{\pi} s(x)\, c(x) \sin 2x\, dx = 0$$

et c'est le cas du n°. 2.

Digitized by Google

On the higher singularities of plane algebraic curves. By Miss C. A. SCOTT. (Communicated by Mr F. P. WHITE.)

[*Received* 15 April, *read* 3 May, 1926.]

Introduction.

1. During the last sixty years the principal questions presented by the higher singularities of plane algebraic curves have been completely solved, and definite results obtained. The two most successful lines of research have been by expansions and quadratic transformation. By each method it has been shown that a higher singularity may be looked upon as containing concealed or "latent" multiple points or lines in addition to those immediately recognized; and from each, with the help of small quantities, has been constructed a topological explanation of these latent multiple elements, which are accounted for as situated in the immediate vicinity of the point and line base of the singularity. Further, by each method it has been proved that as regards the numerical relations known as Plücker's equations a singularity produces the same effect as a definite number of nodes, cusps, bitangents, and stationary tangents.

In the treatment by quadratic transformation as presented by Noether in vol. 23 of the *Math. Annalen* the formal algebraic proofs and topological deductions are so combined that the argument appears to rely on small quantities, whereas they are not really needed except for the topological application. The exposition would gain both in clearness and simplicity if the formal argument were freed from this apparent dependence on small quantities.

The treatment by expansions, on the other hand, is definitely based on the ideas of small quantities, power series, convergence. Now while there can be no question that the concept of small variation underlies most topological discussions, it seems somewhat incongruous in the formal algebraic proofs of properties of the curve as a whole. And while there is no obligation to reject the aid of some special algebraic process simply because it has no geometrical equivalent, yet the fundamental concepts of power series and convergence are in so different a category from those of algebraic geometry that a geometer may be allowed some reluctance to adopt them unless absolutely necessary. The object of the present paper is to show that the procedure in the application of expansions to the proofs of Plücker's equations can be modified so as to dispense entirely with small quantities, keeping the whole question in the domain of finite algebra. As the proofs

Digitized by Google

depend on the intersections of the curve with its polars and the Hessian, it is necessary in the first place to consider the number of points of intersection of two curves that lie at a common point O, that is, the number of O-solutions of two equations

$$F(x, y) = 0, \quad f(x, y) = 0.$$

In the first section it is proved by examination of the x-resultant of F, f that this number depends only on the coefficients of the lower groups of terms in F, f. In the second section this result is applied to show that in all questions relating to the number of O-intersections of two curves each may be replaced by a factored form. In the third section this factored form is used for the proof of Plücker's equations. Of necessity some well-known work has been reproduced in order to make it clear that the modified proofs are really of the character claimed for them, entirely free from dependence on small quantities. This has been given as briefly as possible.

I. *On the O-intersections of two curves.*

2. If
$$F = a_0 z^m + a_1 z^{m-1} + \ldots + a_m,$$
$$f = b_0 z^p + b_1 z^{p-1} + \ldots + b_p,$$

where the a's, b's are homogeneous in x, y, elimination of z from $F = 0, f = 0$ gives an equation $R = 0$, that is, $(x, y)^{mp} = 0$, unless the lowest a, b present are a_r, b_s, when the equation is

$$(x, y)^{mp-rs} = 0.$$

The interpretation of this is that there are $mp - rs$ intersections in addition to the rs immediately detected at O; lines connecting these with O are given by $R = 0$*. But although these points determine definite lines through O, it cannot be taken for granted that they are all distinct from O; as is well known, the elimination of z does not enable us to identify all the O-intersections without more minute investigation.

When small quantities are admitted we formulate the explanation that there may be intersections "in the vicinity of O" in addition to the rs at O. The theory of expansions shows that the detection of one O-intersection in addition to the rs depends on the terms a_r, b_s, the following terms being neglected in comparison with those retained; a second such intersection may depend on a_{r+1}, b_{s+1}, or on a second condition on a_r, b_s, and so on. In

* The vertices of the triangle of reference opposite to the sides x, y, z are distinguished as O_1, O_2, O_3, of which the last is O when non-homogeneous coordinates x, y are used. It is convenient to choose the triangle so that O_1, O_2 are not common to F, f; also if F, f have a common tangent at O_3, this should not be used as a line of reference.

Digitized by Google

general, q intersections in the vicinity of O, in addition to the rs at O, impose q conditions on $a_r, a_{r+1}, \ldots$; $b_s, b_{s+1}, \ldots$, involving not more than the first q of each of the sets a, b, namely

$$K_1(a_r;\ b_s) = 0,$$
$$K_2(a_r, a_{r+1};\ b_s, b_{s+1}) = 0,$$
$$\vdots \qquad \vdots \qquad \vdots$$
$$K_q(a_r, a_{r+1}, \ldots a_{r+q-1};\ b_s, b_{s+1}, b_{s+q-1}) = 0.$$

In these it is not necessary that all the a's and b's indicated be present, but none are present unless expressed.

3. If small quantities are not admitted, this argument is no longer available, a different procedure is needed. The distinction can no longer be drawn between intersections *at* O and intersections *in the vicinity of* O, we can recognize only O-intersections of the curves F, f, or rather O-solutions of the equations $F = 0, f = 0$. It is necessary therefore to prove, by finite algebraic processes, that the conditions for $rs + q$ O-solutions do not involve a_{r+h}, b_{s+h}, if $h \geqslant q$. In other words, it is to be shown that the conditions for q O-solutions beyond the rs at once recognized from the form of the equations involve at most the first q a's and the first q b's.

The process here used is to rearrange the given F, f by powers of x, and form the x-resultant R_x. The equation $R_x = 0$, i.e. $(y, z)^{mp} = 0$, gives mp lines from O_1 to the mp common points. If q of these points are at O_s, then R contains y^q as a factor; the necessary conditions are that the coefficients of $y^0, y^1, \ldots y^{q-1}$ shall vanish. But also y will be a factor if an intersection of F, f is on $y = 0$, even though not at O_s—a repeated factor if two or more intersections lie on $y = 0$, or again if a point, a single intersection of $y = 0$ with F and f, not at O_s, is a point of contact of F, f; so the conditions thus found will involve irrelevant factors. To remove these we may take into account also the y-resultant R_y, and consider a factor x^q in this. The necessary and sufficient conditions for q O-intersections will present themselves in both cases; the obviously irrelevant conditions must be rejected, and the relevant conditions looked for in the common part.

Ex. The two cubic expressions

$$F = \alpha_{10}x + \alpha_{01}y + \alpha_2 + \alpha_3,$$
$$f = \beta_{10}x + \beta_{01}y + b_2 + b_3$$

may be replaced by

$$F = \alpha_{10}x + \alpha_{01}y + \alpha_{20}x^2 + \alpha_{11}xy + \alpha_{02}y^2 + \alpha_{30}x^3 + \alpha_{21}x^2y + \alpha_{12}xy^2 + \alpha_{03}y^3,$$
$$\phi = \qquad \delta y + \delta_{20}x^2 + \delta_{11}xy + \delta_{02}y^2 + \delta_{30}x^3 + \delta_{21}x^2y + \delta_{12}xy^2 + \delta_{03}y^3,$$

where $\delta = \delta_{01} = \begin{vmatrix} \alpha_{10} & \alpha_{01} \\ \beta_{10} & \beta_{01} \end{vmatrix}$, $\delta_{\mu\nu} = \begin{vmatrix} \alpha_{10} & \alpha_{\mu\nu} \\ \beta_{10} & \beta_{\mu\nu} \end{vmatrix}$.

Digitized by Google

For these forms the x-resultant is found to be

$$R_x = A\delta Dy - (\Delta D + B\delta D - \delta\Theta)\, y^2 + \ldots + y^9,$$

where

$$A = -a^2_{10}, \quad B = 2\,(a_{10}a_{11} - a_{01}a_{20}),$$

$$\Delta = \begin{vmatrix} \delta_{20} & \delta_{11} & \delta_{02} \\ a_{10} & a_{01} & . \\ . & a_{10} & a_{01} \end{vmatrix}, \qquad D = \begin{vmatrix} a_{10} & a_{20} & a_{30} \\ \delta_{20} & \delta_{30} & . \\ . & \delta_{20} & \delta_{30} \end{vmatrix},$$

$$\begin{aligned}\Theta = \delta\,\{&(a^2_{30} - 3a_{10}a_{30})\,a_{20}\delta_{30} - (a^2_{20} - 2a_{10}a_{30})\,a_{30}\delta_{20}\} + a_{01}a_{10}\delta_{30}\,(a_{20}\delta_{30} - a_{30}\delta_{20}) \\ &- \delta_{11}a_{10}\,\{(a^2_{20} - 2a_{10}a_{30})\,\delta_{30} - a_{20}a_{30}\delta_{20}\} - a_{11}a^2_{10}\delta^2_{30} \\ &+ \delta_{21}a^2_{10}\,(a_{20}\delta_{30} - 2a_{30}\delta_{20}) + a_{21}a^2_{10}\delta_{20}\delta_{30}.\end{aligned}$$

The (x, y) symmetry in δ, Δ, and the lack of this symmetry in A, B, D, Θ show that together with

$$R_x = A\delta Dy - (\Delta D + B\delta D - \delta\Theta)\, y^2 + \ldots + y^9,$$

we may consider also

$$R_y = A'\delta D'x - (\Delta D' + B'\delta D' - \delta\Theta')\, x^2 + \ldots + x^9.$$

The condition for one extra O-intersection is $A\delta D = 0$; it is also $A'\delta D' = 0$; it is therefore $\delta = 0$. There will be a second extra O-intersection if further

$$\Delta D + B\delta D - \delta\Theta = 0 \quad \text{and} \quad \Delta D' + B'\delta D' - \delta\Theta' = 0,$$

that is, on account of $\delta = 0$, if $\Delta D = 0$ and $\Delta D' = 0$, so the condition is $\Delta = 0$.

As to the remaining factors: if $D = 0$ one intersection is on $y = 0$, not at O; and if also $\Theta = 0$, F, f have contact at this point, unless the vanishing of Θ is caused by $\delta_{20} = 0$, $\delta_{30} = 0$, which in virtue of $D = 0$ are equivalent to one new condition only, the condition for a third intersection on $y = 0$.

4. For the present purpose however it is not necessary actually to resolve any set of conditions into relevant and irrelevant factors; the desired result may be obtained more simply from the principle that if any one coefficient in a_h is necessarily excluded from the conditions for a certain number of O-solutions, then all the coefficients in a_h are excluded. For the linear transformation $x = x' + ky'$, $y = lx' + y'$, where k, l are arbitrary, leaving O unchanged, does not affect the number of O-solutions, though on account of the change of the lines of reference x, y it may modify the irrelevant factors. Now this transformation changes a_h, that is

$$a_{h,0}x^h + a_{h-1,1}x^{h-1}y + \ldots + a_{0,h}y^h$$

into

$$a'_{h,0}x'^h + a'_{h-1,1}x'^{h-1}y' + \ldots + a'_{0,h}y'^h,$$

where any a' is a linear function of all the coefficients $(a)_h$. So if any one of the coefficients $(a)_h$ is present, then for arbitrary values of k, l all the coefficients $(a')_h$ appear in the corresponding condition given by the transformed equation; or, in the more convenient form first enunciated: if any one coefficient in a_h is excluded, all are excluded.

Digitized by Google

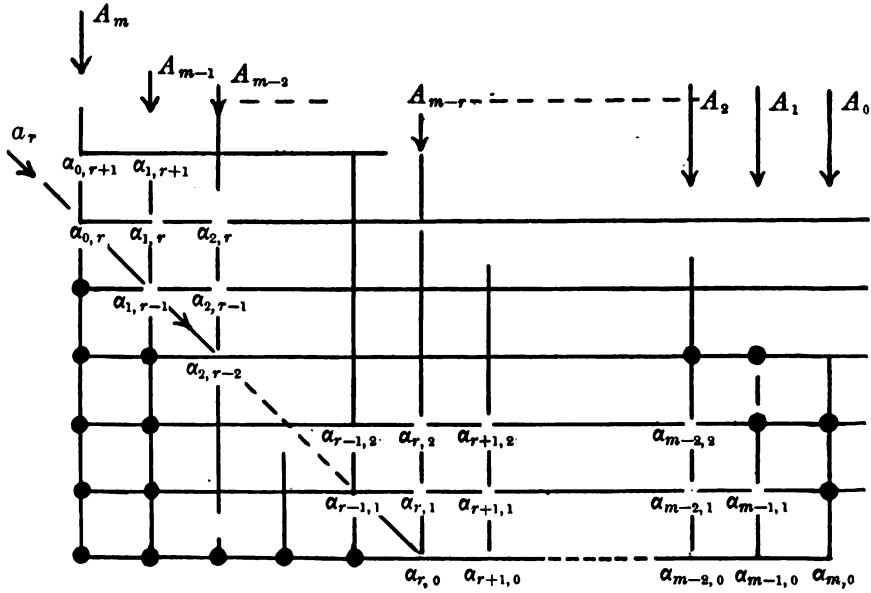
5. In order to form the x-resultant of F, f it is necessary to write each expression by powers of x, so

$$F = a_r z^{m-r} + a_{r+1} z^{m-r-1} + \ldots + a_m,$$
$$= \Sigma a_{\mu,\nu} x^\mu y^\nu \text{ in the non-homogeneous form,}$$

becomes

$$F = A_m + A_{m-1} x + \ldots + A_0 x^m,$$

in which the A's are homogeneous expressions in y, z, with coefficients as indicated in the accompanying scheme.



Thus

$$A_m = a_{0,r} y^r z^{m-r} + a_{0,r+1} y^{r+1} z^{m-r-1} + \ldots + a_{0,m} y^m = y^r \mathbf{A}_r,$$
$$A_{m-1} = a_{1,r-1} y^{r-1} z^{m-r} + \ldots\ldots\ldots\ldots + a_{1,m-1} y^{m-1} = y^{r-1} \mathbf{A}_{r-1},$$
$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$
$$A_{m-r+\mu} = a_{r-\mu,\mu} y^\mu z^{m-r} + \ldots\ldots\ldots + a_{r-\mu,m-r+\mu} y^{m-r+\mu} = y^\mu \mathbf{A}_\mu,$$
$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$
$$A_{m-r+1} = a_{r-1,1} y z^{m-r} + \ldots\ldots\ldots\ldots + a_{r-1,m-r+1} y^{m-r+1} = y \mathbf{A}_1,$$

while

$$A_{m-r} = a_{r,0} z^{m-r} + a_{r,1} y z^{m-r-1} + \ldots + a_{r,m-r} y^{m-r},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$A_\mu = a_{m-\mu,0} z^\mu + a_{m-\mu,1} y z^{\mu-1} + \ldots + a_{m-\mu,\mu} y^\mu,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$A_1 = a_{m-1,0} z + a_{m-1,1} y,$$
$$A_0 = a_{m,0};$$

Digitized by Google

with similar expressions for the coefficients in f, viz:

$$\begin{aligned}
B_p &= y^s \mathbf{B}_s, \\
B_{p-1} &= y^{s-1} \mathbf{B}_{s-1}, \\
&\vdots \\
B_{p-s+1} &= y \mathbf{B}_1, \\
B_{p-s} &= \beta_{s,0} z^{p-s} + \beta_{s,1} y z^{p-s-1} + \ldots + \beta_{s,p-s} y^{p-s}, \\
&\vdots \\
B_1 &= \beta_{p-1,0} z + \beta_{p-1,1} y, \\
B_0 &= \beta_{p,0}.
\end{aligned}$$

The x-resultant of F, f is

$$R = \left| \begin{array}{llllll}
A_m & A_{m-1} & A_{m-2} \ldots & A_0 & . & \ldots \\
. & A_m & A_{m-1} \ldots & A_1 & A_0 & \ldots \\
\multicolumn{6}{l}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots} \\
\multicolumn{6}{l}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots} \\
B_p & B_{p-1} & B_{p-2} \ldots & B_0 & . & \ldots \\
. & B_p & B_{p-1} \ldots & B_1 & B_0 & \ldots \\
\multicolumn{6}{l}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots}
\end{array} \right| \begin{array}{l} p \text{ rows of } A\text{'s} \\ \\ \\ \\ m \text{ rows of } B\text{'s}; \\ \\ \end{array}$$

hence

$$R = \left| \begin{array}{llllllll}
y^r \mathbf{A}_r & y^{r-1} \mathbf{A}_{r-1} \ldots & y \mathbf{A}_1 & A_{m-r} \ldots & A_0 & . & \ldots \\
. & y^r \mathbf{A}_r \quad \ldots & . & y \mathbf{A}_1 & \ldots & A_1 & A_0 & \ldots \\
\multicolumn{8}{l}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots} \\
y^s \mathbf{B}_s & y^{s-1} \mathbf{B}_{s-1} \ldots & y \mathbf{B}_1 & B_{p-s} & \ldots & B_0 & . & \ldots \\
. & y^s \mathbf{B}_s \quad \ldots & . & y \mathbf{B}_1 & \ldots & B_1 & B_0 & \ldots \\
\multicolumn{8}{l}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots}
\end{array} \right| \begin{array}{l} p \\ \text{rows} \\ \\ \\ m \\ \text{rows.} \\ \\ \end{array}$$

If we now manipulate this determinant in the usual way, that is, multiply the first s of the p A-rows by $y^s, y^{s-1}, \ldots y^2, y$, and the first r of the m B-rows by $y^r, y^{r-1}, \ldots y$, so introducing the factor y with exponent

$$\tfrac{1}{2} s(s+1) + \tfrac{1}{2} r(r+1),$$

we put in evidence in the first $r+s$ columns the factor y with exponents

$$r+s, r+s-1, \ldots 2, 1,$$

that is, y with total exponent $\frac{1}{2}(r+s)(r+s+1)$. Thus we see that R contains as a factor y with the exponent

$$\tfrac{1}{2}(r+s)(r+s+1) - \tfrac{1}{2} s(s+1) - \tfrac{1}{2} r(r+1) = rs.$$

Digitized by Google

Hence $R = y^{rs}R'$, where

$$R' = \left|\begin{array}{cccccc} \mathbf{A}_r & \mathbf{A}_{r-1} \ldots & \mathbf{A}_1 & A_{m-r} & yA_{m-r-1} & y^2A_{m-r-2}\ldots \\ . & \mathbf{A}_r \ldots & \mathbf{A}_2 & \mathbf{A}_1 & A_{m-r} & yA_{m-r-1}\ldots \\ \ldots\ldots & & & & & \\ \ldots\ldots & & & & & A_0 \\ \ldots\ldots & & & & & \\ \mathbf{B}_s & \mathbf{B}_{s-1} \ldots \mathbf{B}_1 & B_{p-s} & yB_{p-s-1} & y^2B_{p-s-2} & \ldots\ldots \\ . & \mathbf{B}_s \ldots \mathbf{B}_2 & \mathbf{B}_1 & B_{p-s} & yB_{p-s-1} & \ldots\ldots \\ \ldots\ldots & & & & & \\ \ldots\ldots & & & & & B_0 \end{array}\right|$$

in which, in the first s rows, all A's after A_{m-r} are multiplied by some power of y, and similarly for the B's.

Now in R'

(i) Since

$$\mathbf{A}_\mu = \alpha_{r-\mu,\mu}z^{m-r} + \alpha_{r-\mu,\mu+1}yz^{m-r-1} + \ldots + \alpha_{r-\mu,m-r+\mu}y^{m-r}, \qquad (\mu = 1, 2, \ldots r)$$

it is seen that the second suffix of any α in $\mathbf{A}_\mu$ exceeds the exponent of the accompanying y by μ, where $\mu \leqslant r$. So the second suffix of any α involved $\leqslant r +$ exponent of y.

(ii) Since

$$A_\mu = \alpha_{m-\mu,0}z^\mu + \alpha_{m-\mu,1}yz^{\mu-1} + \ldots + \alpha_{m-\mu,\mu}y^\mu,$$

the second suffix of any α in A_μ is equal to the exponent of y, and therefore $< r +$ exponent of y. Still more in any $y^\lambda A_\mu$ the second suffix $< r +$ exponent of y.

Hence if any contribution to the term y^q in the development of R' is built up as y^{q_1} due to A's and $\mathbf{A}$'s, y^{q_2} due to B's and $\mathbf{B}$'s, where $q_1 + q_2 = q$, every second α-suffix $\leqslant r + q_1$, therefore $\leqslant r + q$, and every second β-suffix $\leqslant s + q_2$, therefore $\leqslant s + q$. This applies to the total coefficient of y^q in R', consequently

$$R' = C_0 + C_1y + C_2y^2 + \ldots + C_qy^q + \ldots,$$

where in $C_0, C_1, \ldots C_q$ the α's have the second suffix

$$\neq r, r+1, \ldots r+q,$$

and the β's $\qquad \neq s, s+1, \ldots s+q.$

6. The application of this to the question of the conditions for q extra O-solutions is immediate. There must be in R' the factor y^q, which requires $C_0 = 0$, $C_1 = 0$, $\ldots$ $C_{q-1} = 0$; the α's involved have the second suffix $< r + q$. In particular, $\alpha_{0,r+q}$ is excluded; consequently by the principle already enunciated no coefficients are present from the set of terms a_{r+q}, and similarly none are present

Digitized by Google

from b_{s+q}. The same conclusion holds for all following a's and b's. Thus we have obtained the result that the conditions for q O-solutions, beyond the rs made evident by the form of F, f, involve at most the first q a's and the first q b's, though it is not necessary that all these be present. It may be noted that this result is independent of m and p.

In considering the O-solutions of $F=0$, $f=0$ we may equally well use $F+lf=0$, $F+l'f=0$. If both F and f begin with terms of degree r, and these lowest terms are equivalent, we may replace f by the combination ϕ (e.g. $F-f$) that removes these terms and possibly others, and instead of

$$F=a_r+a_{r+1}+\ldots,$$
$$f=b_r+b_{r+1}+\ldots,$$

consider

$$F=a_r+a_{r+1}+\ldots,$$
$$\phi=c_s+c_{s+1}+\ldots,$$

where $s>r$.

The results as to the number of O-solutions are necessarily identical, but the form in which these present themselves is slightly different. An illustration will make this clear. With $r=2$, consider the conditions for 7 O-solutions; these depend on not more than $7-2.2$, i.e. 3 a's and 3 b's. The three conditions for the three extra O-solutions are

$$(a_2; b_2)=0, \quad (a_2, a_3; b_2, b_3)=0, \quad (a_2, a_3, a_4; b_2, b_3, b_4)=0,$$

where any of the a's or b's may be absent.

If now $a_2 \equiv b_2$, we replace f by $\phi \equiv F-f$, and consider

$$F=a_2+a_3+a_4+\ldots,$$
$$\phi= \quad c_3+c_4+\ldots.$$

But as now there is only one extra solution $(7-2.3)$ to provide for, we need only one a and one c; there is one condition $(a_2; c_3)=0$. Since however $a_2 \equiv b_2$, and $c_3=a_3-b_3$, this is $(a_2, a_3; b_2, b_3)=0$.

The explanation is, of course, that the stated equivalence of a_2, b_2 has imposed two conditions, so the three to be expected in the general case $(a_2 \not\equiv b_2)$ reduce to

$$(a_2; b_2)_1=0, \quad (a_2; b_2)_2=0, \quad (a_2, a_3; b_2, b_3)=0,$$

a reduction in extent allowed for by the formulation that *not more than* a certain number of a's and b's are involved.

7. The result obtained shows that in considering the possibility of q additional O-solutions of $F=0, f=0$ where

$$F=a_r+a_{r+1}+\ldots,$$
$$f=b_s+b_{s+1}+\ldots,$$

or in assigning the conditions for such extra O-solutions, we may ignore all the a's and b's except the first q of each, not on the ground that the terms thus ignored are small in comparison with those retained and therefore negligible, but simply because they do not enter into the algebraic conditions.

Digitized by Google

If then F_m, $F'_{m'}$ agree as to the lower terms a_r, a_{r+1}, ..., and f_p, $f'_{p'}$ as to b_s, b_{s+1}, ..., and the agreement in each case extends sufficiently far, $F'=0$ and $f'=0$ have the same number of O-solutions as $F=0$ and $f=0$. Consequently in any investigation that is concerned only with the *number* of O-solutions, irrespective of any geometrical explanation of the location of the common points of the curves F, f whether at O or in the vicinity of O, F', f' may be used instead of F, f.

Note. It is convenient to have symbols for the intersections of two curves F, f, or for the solutions of two equations $F=0, f=0$. Those that will now be used are

$\{F, f\}$ for the set of solutions or intersections,

$\{F, f\}_0$ for the set of O-solutions or O-intersections,

$I_0\{F, f\}$ for the *number* of O-solutions or O-intersections.

II. *Factoring of* $F(x, y)$.

8. It will now be shown that in certain investigations $F(x, y)$ may be replaced by a factored form. The proof applies not only when the exponents of x, y in F are integral, but also when the exponents of x are integral or fractional, provided always the exponents of y are integral. For simplicity of treatment and generality of statement the lines of reference $x=0$, $y=0$ are chosen so that no power of x is a factor in the lowest terms a_k. It follows that these terms include y^k, hence k is an integer, and the exponents of x in a_k are all integral.

Let $\tau \equiv y - f(x)$, where $f(x) = \Sigma A_i x^{\lambda_i}$, the λ's being integral or fractional $(1 \leqslant \lambda_1 < \lambda_2 < \ldots \leqslant \sigma)$, then by actual division

$$F(x, y) \equiv \tau F'(x, y) + X(x),$$

irrespective of any particular choice of $f(x)$. In this identity

$$X(x) \equiv F\{x, f(x)\} = \Sigma g_\mu x^\mu.$$

Since $f(x)$ contains no higher power of x than x^σ, and F no higher power of y than y^m, the greatest exponent of x in $X(x)$ is not greater than $m\sigma$, so $X(x)$ consists of a finite number of terms $g_\mu x^\mu$, where μ has a finite range. The lowest exponent depends on $f(x)$, which will now be chosen so that this remainder, X, shall have a certain characteristic.

If there is an expression $\Sigma A_i x^{\lambda_i} (1 \leqslant \lambda_1 < \lambda_2 < \ldots \leqslant \sigma)$ which substituted for y in

$$F \equiv a_k + a_{k'} + \ldots \quad (k < k' < k'' < \ldots)$$

Digitized by Google

gives as a result $\phi(x)$, in which the lowest term is x^π, where π is integral or fractional, let this be used as $f(x)$. The substitution $y = f(x)$, i.e. $\tau = 0$, in the identity

$$F \equiv \tau F' + X(x)$$

gives the result $X(x)$. Hence this remainder, $X(x)$, is precisely $\phi(x)$, in which the lowest term is x^π, and the product $\tau F'$ agrees with F as to all terms of degree less than π. Consequently the product $\tau F'$ may be used instead of F in any investigation that involves no a of suffix π or higher. If then it is found possible to make π sufficiently great, we shall be able to substitute $\tau F'$ for F in all questions relating to O-intersections of F with other curves.

Comparison of the lowest terms on the two sides of the identity

$$F \equiv \tau F' + X,$$

where

$$\tau = y - A_1 x - \Sigma A_\lambda x^\lambda,$$

$$F' = v_h + v_{h'} + \ldots, \quad (h < h' < h'' < \ldots)$$

shows

$$(y - A_1 x)\, v_h \equiv a_k,$$

hence $y - A_1 x$ is a factor in a_k, and v_h is the quotient, so $h = k - 1$.

It is seen that in F', as in F, the lowest terms have only integral exponents, and contain no power of x as a factor.

9. The conclusion as to the character of the remainder X depends on the existence of a finite number of terms of an expression $f(x)$ satisfying certain conditions with respect to $F(x, y)$. That such an expression—*an expansion for* y—exists, provided x is not a factor in a_k, is shown by the usual procedure for finding the terms in succession, with ascending exponents, arranged in the form that is theoretically simplest, depending on the direct application of the defining condition that the lower powers of x in $F(x, f)$ shall vanish. For the present purpose it is not necessary to examine minutely the character of the expansion; if alternative values offer themselves for the exponent of the term sought or for its coefficient it suffices to use any one of these, as all that is needed is the assurance that some one expansion can be found. This assurance can be obtained very simply.

If we have already found

$$y = f(x) = A_1 x + \ldots + A_\sigma x^\sigma,$$

which reduces $F(x, y)$ to X, that is to $g_0 x^\pi +$ higher terms where $\pi > \sigma$, and no modification of the A's or insertion of other terms of exponent $< \sigma$ will increase π, it is to be shown that we can find another term, $\eta, = A x^\theta$, where $\theta > \sigma$, such that $y = f(x) + \eta$ shall reduce F to X', in which the lowest exponent $\pi' > \pi$. For this the finite algebraic development

$$F(x, y + \eta) = F(x, y) + \eta F^{(1)} + \ldots + \frac{1}{h!}\eta^h F^{(h)} + \ldots$$

Digitized by Google

suffices. Since k is integral, and a_k contains a term y^k,

$$F^{(k)} = g_k + \text{terms in } x,$$

while $F^{(h)}$, $h \neq k$, may begin with a power of x, exponent π_h. So

$$F(x, y+\eta) = x^{\pi}(g_0 + X_0) + \eta x^{\pi_1}(g_1 + X_1) + \tfrac{1}{2}\eta^2 x^{\pi_2}(g_2 + X_2)$$
$$+ \ldots + \frac{1}{k!}\eta^k (g_k + X_k) + \ldots \frac{1}{h'!}\eta^{h'} x^{\pi_{h'}}(g_{h'} + X_{h'}) + \ldots.$$

Since $\eta = Ax^{\theta}$, the lowest exponents in the successive terms are

$$\pi, \pi_1 + \theta, \ldots, \pi_h + h\theta, \ldots, k\theta, \ldots, \pi_{h'} + h'\theta, \ldots.$$

It is a simple matter to show that if we take for θ the greatest of such of the quanties $\frac{1}{\mu}(\pi - \pi_\mu)$ as are positive, certain of these exponents, for values $\mu, \mu', \mu'', \ldots$, become equal to π, and the rest all greater than π. It is worth noting that if $h' > k$, $\pi_{h'} + h'\theta > k\theta$ so $> \pi$, hence it is not necessary to carry the development of $F(x, y+\eta)$ beyond the term $\frac{1}{k!}\eta^k F^{(k)}$. In this way θ is determined linearly *; A is obtained by equating to zero the complete coefficient of x^{π}, hence it is any root of the equation

$$g_0 + g_\mu A^\mu + g_{\mu'} A^{\mu'} + \ldots = 0$$

in which $g_0 \neq 0$. There will be alternative or possibly repeated values for A unless $\mu = 1$, $\mu' = \mu'' = \ldots = 0$, but this is immaterial as only one solution is needed. Since A is thus determined, we cannot control the coefficients of π' etc.; further, since

$$\pi_\mu + \mu\theta = \pi,$$

it follows that $\theta \leqslant \pi$, hence $\pi' > \theta$. So the expansion

$$y = A_1 x + \ldots + A_\sigma x^\sigma + Ax^\theta$$

has the characteristics specified for

$$y = A_1 x + \ldots + A_\sigma x^\sigma.$$

Hence the process may be continued indefinitely, but this is not needed here, as it will appear that the expansion is to be carried only to some predetermined exponent σ; the procedure shows that the lowest exponent in the remainder X is then greater than σ.

* This process, which gives at every stage the *greatest* admissible value for θ, may possibly be the same as that of John Stewart of Aberdeen (1745) mentioned by de Morgan: "Newton's Method of Coordinated Exponents," *Trans. Camb. Phil. Soc.* IX, 1856.

Digitized by Google

Ex. 1. If $F=(y-x)^3-xy(y-x)+x^3(y-x)+x^2y^3+x^6$,
the process selects $y=x+x^3+\ldots$;
then $F'=y-x-xy+2x^3+\ldots.$
The other expansion is $y=x+x^2+\ldots.$

Ex. 2. If
$$F=(y-x)^3+y^4-3xy^3+2x^2y^2-2x^3y^2+3x^4y+x^5y+x^7,$$
the process selects $y=x+x^2+\ldots$;
then $F'=(y-x)^2+y^3-2xy^2+x^2y-x^3+\ldots.$

The remaining expansions are $y=x\pm x^{\frac{3}{2}}+\ldots.$

10. More precise information as to the relation of σ to π, the lowest exponent in X, may be obtained from the identity
$$F\equiv\tau F'+X.$$

If the exponents in F and in τ so far as used, therefore also in F', are integral, the product
$$(y-A_1x-A_2x^2-\ldots-A_\sigma x^\tau)(v_{k-1}+v_k+\ldots+v_\rho)$$
agrees with F up to and including terms built up as
$$(y-A_1x)v_\rho-A_2x^2v_{\rho-1}-\ldots-A_\sigma x^\sigma v_{k-1};$$
hence $\rho+1=\sigma+k-1=\pi-1,$
that is $\rho=\sigma+k-2,\ \pi=\sigma+k,$
and the product
$$(y-A_1x-A_2x^2-\ldots-A_\sigma x^\sigma)(v_{k-1}+\ldots+v_{\sigma+k-2})$$
agrees with F up to and including the terms of degree $\sigma+k-1$. So $F\equiv\tau F'+X$, where the lowest terms in X are of degree $\sigma+k$.

If there are fractional exponents in F or τ, let $\sigma=\frac{\lambda}{q}$; the agreement with F extends over terms of degree $\frac{\lambda}{q}+k-1$, terms potentially present in F with zero coefficients if necessary. It suffices however to state that the agreement covers terms of a certain integral degree, the integer next smaller than the correct fractional value; but as a matter of fact it is not necessary in either case to state the actual extent of agreement, we can make this as high as we please by choosing for σ a sufficiently great finite value.

Note. As an alternative the remainder $X(x)$ may be obtained as $X(x, y)$, either by leaving the division incomplete or by substituting for some power of x the equivalent given by the expression for τ. It may be noted that although in the proof given it is taken for granted that for the division by τ, F is arranged by descending powers of y, so that the remainder appears as $X(x)$, yet when we know the desired form of result, the division by τ, that is by $(y-A_1x)-A_\lambda x^\lambda-\ldots-A_\sigma x^\sigma$, is more conveniently performed when F is arranged in the usual way as $a_k+a_{k'}+\ldots$. The division can then be broken off when the lowest terms in the remainder $X(x, y)$ are of the desired degree.

Digitized by Google

11. Since F' is obtained from F as the quotient, with

$$y - A_1 x - \Sigma A_\lambda x^\lambda$$

as divisor, all the y-exponents in F' are integral, the x-exponents may be integral or fractional, and as already pointed out the lowest terms are integral and contain no power of x as a factor. Hence we can apply to F' the same process, obtaining the identity

$$F' \equiv \tau' F'' + X',$$

where τ' is an expansion for F'. From this

$$\begin{aligned} F &\equiv \tau(\tau' F'' + X') + X \\ &\equiv \tau\tau' F'' + (\tau X' + X) \\ &\equiv \tau\tau' F'' + X''(x, y), \end{aligned}$$

where the lowest terms in X'' can, if desired, be made of the same degree π as those in X by making those in X' of degree $\pi - 1$ or higher.

This identity, written as

$$F \equiv \tau' G + Y,$$

shows that τ', obtained as a terminated expansion for F', is also a terminated expansion for F. It is true conversely that τ', a terminated expansion for F, *if different from* τ (a condition not necessary in the direct theorem) is a terminated expansion for F'. Let τ, τ' differ at x^λ, and let $\tau' = 0$ reduce F to X', where the lowest terms are of degree π. Substitution of $\tau' = 0$ in the identity

$$F \equiv \tau F' + X$$

shows that the lowest terms in $\tau F'$ are made to be of degree $\geqslant \pi$, so the lowest terms in F' are of degree $\geqslant \pi - \lambda$, which is at our disposal on account of π. Hence τ' is a terminated expansion for F'.

By repetition of the factoring we have

$$\begin{aligned} F &\equiv \tau_1 F_1 + X_1 \\ F_1 &\equiv \tau_2 F_2 + X_2 \\ F_2 &\equiv \tau_3 F_3 + X_3 \text{ etc.} \\ &\vdots \quad\quad \vdots \quad\quad \vdots \end{aligned}$$

from which

$$\begin{aligned} F &\equiv \tau_1 \tau_2 \ldots \tau_h F_h + X_h \\ &\equiv \mathbf{T}_h F_h + X_h, \end{aligned}$$

where $\mathbf{T}_h$ is the product of h terminated expansions for F. It is immaterial whether these are different or identical. If all are carried to the same exponent σ (conveniently stated as integral, with zero coefficient if necessary), the identity shows that the

Digitized by Google

lowest terms in F_h are of degree $k-h$, the highest of degree $k-h+\sigma-1$; the lowest terms in X_h are of degree greater than $k+\sigma-1$. Since F_h and X_h are the quotient and remainder when F is divided by the product $\mathbf{T}_h$, the order in which the factors $\tau_1, \tau_2, \ldots \tau_h$ are introduced makes no difference.

So long as h is less than k the factoring process can be continued; hence there are at least k expansions $\tau_1, \tau_2, \ldots \tau_k$, different or (apparently) the same. It will be shown later that if carried far enough these are all different, though if the lowest terms in F contain a repeated factor, l-fold, l of the τ's agree to some extent. That there is no expansion different from the k thus individually segregated in the course of the factoring is a consequence of the theorem proved above, that an expansion for F, different from τ_1, is an expansion for F_1; consequently if different from τ_2, it is an expansion for F_2, and so on through the k expansions $\tau_1, \tau_2, \ldots \tau_k$.

In the identity

$$F \equiv \mathbf{T}F_k + X_k$$

the final quotient F_k, which will now be called $\mathbf{N}$, begins with an absolute term u_0, different from zero. It is convenient to write F so that this shall be unity, then

$$F \equiv \mathbf{NT} + X(x, y),$$

where $\mathbf{T}$ is the product of k factors

$$y - A_1x - \Sigma A_\lambda x^\lambda - A_\sigma x^\sigma,$$

$\mathbf{N}$ is

$$1 + u_{\lambda_1} + u_{\lambda_2} + \ldots + u_{\sigma-1},$$

and the lowest terms in X are of degree $> k+\sigma-1$, the highest $\not> k\sigma+\sigma-1$ or m, which ever may be the greater.

12. If in one expansion τ_1 there are fractional exponents with L.C.D. $= q$, then as $\tau_1 = 0$, that is $y = \phi(x_1^{1/q})$, substituted in

$$F \equiv \tau_1 F_1 + X_1$$

makes

$$F(x, y) = X_1 = \rho_0 + \rho_1 x_1^{1/q} + \rho_2 x_1^{2/q} + \ldots,$$

where the ρ's involve only integral powers of x, it follows that $\tau_2 = 0$, that is $y = \phi(x_2^{1/q})$, the corresponding expansion in terms of a second value of $x^{1/q}$, will make

$$F(x, y) = X_2 = \rho_0 + \rho_1 x_2^{1/q} + \rho_2 x_2^{2/q} + \ldots,$$

where the ρ's are the same as before. Hence the lowest terms in X_2 are of the same degree as those in X_1; τ_2 is an expansion for F, the remainders X_1, X_2 are conjugate as also the quotients F_1, F_2. The terminated expansion τ_1 is thus seen to be one of a set of q conjugate terminated expansions. In the product $\mathbf{T}$ these may be multiplied together, giving the rational product Q, the terminated algebraic representation of a superlinear branch.

Digitized by Google

We thus arrive at the finite rational algebraic identity

$$F \equiv \mathbf{N} \cdot \tau_1 \tau_2 \dots QQ' \dots + X(x, y)$$

as an alternative to

$$F \equiv \mathbf{N} \cdot \tau_1 \tau_2 \dots \tau_k + X(x, y),$$

in which all k expansions, rational or irrational, are explicitly present.

These investigations may be summed up in the theorem:

If the lowest terms in $F(x, y)$, of degree k, do not contain x as a factor, we can find in one way only a finite product $\mathbf{N} \cdot \tau_1 \tau_2 \dots \tau_k$, the resolved form of F, in which any τ is of the form

$$y + A_1 x + \Sigma_\sigma A_\lambda x^\lambda;$$

this resolved form differs from F only by a finite number of terms of degree not less than a finite quantity π at our disposal, and may be substituted for F in any discussion that involves no terms of F of degree as great as π.

In the proof of this theorem no use has been made of infinite series, and there is no dependence on small quantities.

It is convenient to call the terminated expansions "elements," and any Q, the rational product of a set of conjugate elements, a "component." Thus a component is the finitely terminated equivalent of a complete branch. A terminated linear expansion is indifferently an element or a component.

13. From the theorem now proved it follows that in determining the number of O-solutions of $F = 0$, $F' = 0$, i.e. $I_0\{F, F'\}$, we may replace each by the resolved form and consider only $I_0\{\mathbf{NT}, \mathbf{N'T'}\}$. Further, on account of the absolute term in $\mathbf{N}$, $\mathbf{N'}$ these factors make no contribution to the O-solutions, so may be discarded; consequently $I_0\{F, F'\} = I_0\{\mathbf{T}, \mathbf{T'}\}$. Elimination of y from $\mathbf{T} = 0$, $\mathbf{T'} = 0$, that is, formation of the y-resultant, gives $R(\mathbf{T}, \mathbf{T'}) = 0$, an equation in x; the exponent of the lowest power of x in this is the number of O-solutions of $\mathbf{T} = 0$, $\mathbf{T'} = 0$, and therefore of $F = 0$, $F' = 0$. Since

$$\mathbf{T} = \tau_1 \tau_2 \dots \tau_h \text{ and } \mathbf{T'} = \tau_1' \tau_2' \dots \tau_k'$$

this resultant may be formed as

$$\Pi_{ij}(\tau_i - \tau_j') \quad \{i = 1, 2, \dots h;\ j = 1, 2, \dots k\},$$

and if the exponent of the lowest power of x in $\tau_i - \tau_j'$, that is, the exponent of the power of x at which τ_i, τ_j' begin to differ, is denoted by $[i, j]$, the exponent of the lowest power of x in $R(\mathbf{T}, \mathbf{T'})$ is $\Sigma[i, j]$. In this any of the separate contributions $[i, j]$ may be fractional, but $\Sigma[i, j]$ is certainly integral, being the exponent of a power of x in the y-resultant of $\mathbf{T}$, $\mathbf{T'}$, also in the y-resultant of F, F', both obtainable rationally.

Digitized by Google

In geometrical language $I_0\{F, F\}$ is the number of O-intersections of F, F'. Cayley and others use the term "number of intersections" also for two partial branches, even though this number be fractional—a verbal convention that seems unnecessarily strained though causing no confusion. When small quantities are admitted "order of evanescence" or an equivalent (Smith, Halphen, etc.) is appropriate, but as this is not available here the term "relative index" is used, defined as the exponent of the power of x at which two elements differ. To avoid unnecessary increase in the number of symbols $I_0\{\tau, \tau'\}$ is used to express "relative index" or "number of O-solutions" or "O-intersections" as the case may be. So

$$I_0\{F, F'\} = I_0\{\mathbf{T}, \mathbf{T}'\} = \Sigma I_0\{\tau_i, \tau_j'\} \begin{Bmatrix} i = 1, 2, \dots h \\ j = 1, 2, \dots k \end{Bmatrix}.$$

14. The resolved form of F may be used not only in determining the number of O-intersections with the polars or the Hessian, but also in forming admissible substitutes for these.

(i) *F and the polars.* Let ΔF be any first polar;

$$F \equiv \mathbf{NT} + X,$$

so

$$\Delta F \equiv \Delta(\mathbf{NT}) + \Delta X$$

by the law of formation of a polar. Hence ΔF differs from $\Delta(\mathbf{NT})$ by terms of degree $\geqslant \pi - 1$. So for a sufficiently great value of π, certainly less than $m(m-1) - k(k-1)$, we may use the resolved form of F.

Further, in forming the polar we may ignore the factor $\mathbf{N}$. For

$$\Delta(\mathbf{NT}) = \mathbf{N} \,.\, \Delta\mathbf{T} + \mathbf{T} \,.\, \Delta\mathbf{N},$$

so the O-solutions of $\mathbf{NT} = 0$ with $\Delta(\mathbf{NT}) = 0$ are

$$\text{(i)}\ \{\mathbf{N}, \Delta\mathbf{N}\}_0 + \{\mathbf{N}, \mathbf{T}\}_0, \qquad \text{(ii)}\ \{\mathbf{T}, \Delta\mathbf{T}\}_0 + \{\mathbf{T}, \mathbf{N}\}_0.$$

But $\mathbf{N}$ has an absolute term, so the only O-solutions are

$$\{\mathbf{T}, \Delta\mathbf{T}\}_0;$$

hence

$$\{\mathbf{NT}, \Delta(\mathbf{NT})\}_0 = \{\mathbf{T}, \Delta\mathbf{T}\}_0$$

and consequently

$$I_0\{F, \Delta F\} = I_0\{\mathbf{T}, \Delta\mathbf{T}\}.$$

(ii) *F and the Hessian.* Consider first $\{F, HF\}_0$, where

$$F \equiv \phi + X,$$

the lowest terms in X being of degree π in x, y; then

$$H(F) = \begin{vmatrix} \phi_{11} + X_{11} & \phi_{12} + X_{12} & \phi_{13} + X_{13} \\ \phi_{21} + X_{21} & \phi_{22} + X_{22} & \phi_{23} + X_{23} \\ \phi_{31} + X_{31} & \phi_{32} + X_{32} & \phi_{33} + X_{33} \end{vmatrix}.$$

Digitized by Google

Examination of the eight determinants, of which the first is $H(\phi)$, shows that

$$H(F) = H(\phi) + \text{terms of degree} \geqslant \pi - 2 \text{ in } x, y;$$

hence $H(F)$ and $H(\phi)$ differ only after terms in x, y of a certain degree, as high as we please. That is,

if $$F \equiv \phi + X$$

$$H(F) \equiv H(\phi) + Y,$$

so with $$F \equiv \mathbf{NT} + X$$

we find $$H(F) \equiv H(\mathbf{NT}) + Y.$$

Consequently

$$I_0\{F, H(F)\} = I_0\{\mathbf{NT}, H(\mathbf{NT})\},$$

which shows that instead of F and H we may use the resolved form and its Hessian.

Further, in forming the Hessian we may ignore the factor $\mathbf{N}$. The known formula for the Hessian of a product $f_1 f_2$, namely

$$H(f_1 f_2) = m_1 f_2^3 h_1 + m_2 f_1^3 h_2 + \Theta f_1 f_2,$$

where $h_1 = H(f_1)$, $h_2 = H(f_2)$, and m_1, m_2 are numerical multipliers, applied to the product $\mathbf{NT}$, shows that the O-solutions of $\mathbf{NT} = 0$ with $H(\mathbf{NT}) = 0$, that is, with

$$m_1 \mathbf{T}^3 H(\mathbf{N}) + m_2 \mathbf{N}^3 H(\mathbf{T}) + \Theta \mathbf{NT} = 0,$$

are (i) $\{\mathbf{N}, \mathbf{T}^3 H(\mathbf{N})\}_0$, (ii) $\{\mathbf{T}, \mathbf{N}^3 H(\mathbf{T})\}_0$.

Inasmuch as $\mathbf{N}$ has an absolute term, the only O-solutions are

$$\{\mathbf{T}, \mathbf{N}^3 H(\mathbf{T})\}_0, \text{ that is } \{\mathbf{T}, H(\mathbf{T})\}_0;$$

consequently

$$I_0\{F, H(F)\} = I_0\{\mathbf{NT}, H(\mathbf{NT})\} = I_0\{\mathbf{T}, H(\mathbf{T})\}.$$

Thus as regards the O-intersections of F with the polars or the Hessian the factor $\mathbf{N}$ produces no effect, it is a neutral factor. On account of its algebraic rôle, which is to combine the O-elements or O-components so as to reproduce F to the desired extent, it may be called the synthetic factor. The recognition of this factor enables us to base certain investigations on the finite algebraic identity

$$F \equiv \mathbf{N} \cdot \tau_1 \tau_2 \ldots \tau_k + \underset{\pi \to}{X}(x, y),$$

and replace F by the factored form $\tau_1 \tau_2 \ldots \tau_k$.

15. In the proof that there are k elements τ there is nothing that requires these to be different; if the expansions are terminated at a sufficiently low exponent, two or more may be identical. It will now be proved that if the expansions are carried far enough, the k elements are all different. This depends on the relation of F to the polars.

Digitized by Google

An irreducible curve F_m and Δ, the polar of an arbitrary point, have $m(m-1)$ intersections of which a definite number n do not lie at the singular points. This number n may be found without the use of the polars, as it is the degree of the reciprocal equation. Consequently $I_0\{F, \Delta\}$ cannot exceed a certain number D,

$$= m(m-1) - n;$$

the exponent of the lowest power of x in the resultant of F, Δ is not greater than D—or, more conveniently, $R(F, \Delta)$ certainly contains a term $Mx^{D'}$, where $D' \leqslant D$.

Let the τ's in the resolved form of F be carried so far as to allow for D intersections with Δ. On the hypothesis that two of these elements are the same,

$$\mathbf{T} = \tau^2 . \tau_1\tau_2 \dots \tau_{k-2},$$

and
$$\Delta\mathbf{T} = \tau^2[\quad] + \tau[\quad] = \tau\psi.$$

Hence the resolved form of $\Delta\mathbf{T}$ presents τ as one factor,

$$\Delta\mathbf{T} = \mathbf{N}'\mathbf{T}' + X',$$

where
$$\mathbf{T}' = \tau . \tau_1'\tau_2' \dots \tau'_{k-2}.$$

Now
$$I_0\{F, \Delta\} = I_0\{\mathbf{T}, \Delta\mathbf{T}\} = I_0\{\mathbf{T}, \mathbf{T}'\},$$

hence in $R(\mathbf{T}, \mathbf{T}')$ there is a term $Mx^{D'}$, where $D' \leqslant D$. But on account of the common factor τ in $\mathbf{T}$, $\mathbf{T}'$ their resultant, $\Pi(\tau_i - \tau_j')$, vanishes identically and there is no term $Mx^{D'}$, a contradiction which disproves the hypothesis that two elements τ are the same. Hence if the expansions at O are carried sufficiently far the k elements are all different*. In what follows it is to be understood that this condition has been complied with, so that in the identity

$$F \equiv \mathbf{N} . \tau_1\tau_2 \dots \tau_k + X$$

the τ's are all different.

Note. Possibly the product $\mathbf{NT}$ should not be considered "the resolved form" unless the elements are all different. But in investigating the relation of two curves F, F' it may not be necessary to carry the expansions so far, while again it may be necessary to carry them even further than will ensure the divergence of all τ's, so it seems more convenient to let the term "resolved form" apply to any admissible $\mathbf{NT}$, with the understanding that the agreement with F is of the necessary extent.

16. This is essentially the same as the reduction due to Brill, based on the theorem of Weierstrass that it is possible, in one way only, to find a reducing factor

$$\theta(x, y) \equiv 1 + \theta_1 + \theta_2 + \dots,$$

* Cayley, *Quart. Journ. Math.* VII. 1865.

Digitized by Google

such that θF shall contain no power of y above y^k, where k is the degree of the lowest terms in F, and this power only in the one term y^k. The neutral or synthetic factor here used, which is Halphen's Q, terminated, is equivalent to the terminated reciprocal of the reducing factor of Weierstrass. Brill shows* that the resulting form

$$y^k + y^{k-1}xP_1(x) + y^{k-2}x^2P_2(x) + \ldots + x^kP_k(x)$$

can be expressed as the product of k factors $y - x\mathbf{P}_i(x)$, where the P's and $\mathbf{P}$'s are power-series. The procedure is sufficiently simple for *different* factors in a_k, but it becomes complicated when a_k contains a *repeated* factor, which is the important case.

The present investigation offers a more elementary treatment in which, owing to the segregation of the individual factors as they arise, the case of a repeated factor demands no special consideration. There is however a more important difference relating to the concepts involved; in the reduction as presented here no use is made of infinite series and no questions of convergence arise. Likewise in the applications now to be made of the finite algebraic identity

$$F \equiv \mathbf{NT} + X$$

there is no reference to small quantities, though the identity itself affords a convenient basis for the introduction of small quantities for topological purposes.

III. *Plücker's Equations.*

17. From the algebraic point of view Plücker's equations are relations connecting the order and class of a curve, given by the degrees of the point-equation and its reciprocal, with certain numbers obtained algebraically without reference to any geometrical significance, even though with a view to conciseness of statement we may choose to employ such geometrical terms as point, line, tangent, etc. These numbers are the summation over the whole curve of certain indices defined for the singularities.

These indices present themselves separately and in combination in the consideration of the O-intersections of F with the polars and the Hessian, and in the formation of the expansions. Now in the determination of the O-intersections, or of the expansions, it has been shown that we may replace F by $F + \lambda\phi$, where ϕ is arbitrary with the lowest terms of sufficiently high degree. In the pencil $F + \lambda\phi = 0$ the only necessary point-singularities lie at the critic centres, and we can choose λ so as to avoid all these with the exception of O. That is, we can replace F by a curve that has the

* *Sitzungsber. der Münch. Akad.* 1891, p. 207; *Math. Annalen*, 1891, xxxix. p. 129; *Vorl. ü. ebene algebraische Kurven u.s.w.* 1925, pp. 91—113.

Digitized by Google

same O-singularity and no other singular points. Hence the indices belong to the singularity absolutely, they depend neither on the order of the curve nor on the possible presence of other singularities. Further, we may use the resolved form, which exhibits as much as desired of the expansions, in forming the polars and the Hessian. We shall find that this practically reduces the problem to the discussion of the components.

The order and class of a singularity at O on a rational component with $y=0$ as tangent are the degrees of the lowest terms in the point-equation and the (reciprocal) line-equation. In geometrical terms k, the order, is the number of points on an arbitrary line through O that are represented by O; and l, the class, is the number of tangents through an arbitrary point on $y=0$ that are represented by $y=0$. The order, k, is therefore the number of elements in the rational component, that is, the number of conjugate expansions in the complete (superlinear) branch. The reciprocal statement as to the class, l, is that it is the number of conjugate expansions in the complete branch, when expressed in terms of line-coordinates*; it can however be found without determining the reciprocal equation, for it will be proved by means of polars that the component and the O-tangent have $k+l$ O-intersections.

The indices are:

(1) The *discriminantal index* D, a certain exponent in the resultant of F and a polar; that is, if the polar is Δ_y, the exponent of the power of x that is a factor in the y-discriminant of F. This index depends on the singularity as a whole, it may be obtained without any preliminary resolution into components, though its value is most easily found from the completely factored form.

(2) The *index* D', *reciprocal* to D.

(3, 4) The *cuspidal index* κ and the *inflexional index* ι, defined for a rational component with a singularity of order k, class l, as $\kappa=k-1$, $\iota=l-1$. For the whole singularity these indices are $\Sigma(k-1)$ and $\Sigma(l-1)$.

It is found convenient to introduce also

(5, 6) the *nodal index*, defined indirectly for a component or for the whole singularity as the number ν for which $D=2\nu+3\kappa$, and the *reciprocal* τ, for which $D'=2\tau+3\iota$. For present purposes these two indices, ν and τ, are not needed, their significance appears however when topological explanations are desired for results obtained by detailed analysis of a singularity.

* In *Proc. Lond. Math. Soc.* VI. 1873–6, H. J. St. Smith obtains the expansions in line-coordinates, and thus finds l, but as the argument depends on small quantities, it is not admissible here. The same comment applies to the proofs of Halphen (*Mém. prés....à l'Ac. des Sc. de Paris*, XXVI. 1874–7) and Stolz (*Math. Ann.* VIII. 1874–5).

Digitized by Google

It is a simple matter of proof that for the elementary singularities these formal algebraic definitions are consistent with the usual definitions.

18. *Relation of F to the polars.* The first two of Plücker's equations are equivalent to definitions of ν and τ; it remains to show how D, or $2\nu + 3\kappa$, may be found without actually forming the y-discriminant. Except for the more compact statement made possible by the use of the resolved form, and the omission of any reference to power-series and convergence, the determination of the value of D is the same as that now current, due to Cayley. It is given here briefly for the sake of continuity of argument.

By definition

$$D = I_0\{F, \Delta_y F\} = I_0\{\mathbf{T}, \Delta_y \mathbf{T}\}.$$

Hence if

$$\mathbf{T} = \tau_1\tau_2 \dots \tau_k = \tau_1 \mathbf{T}_1 = \tau_2 \mathbf{T}_2 = \dots,$$

so that

$$\begin{aligned}\Delta_y \mathbf{T} &= \mathbf{T}_1 \Delta\tau_1 + \mathbf{T}_2 \Delta\tau_2 + \dots \\ &= \mathbf{T}_1 + \mathbf{T}_2 + \dots + \mathbf{T}_k, \text{ since } \Delta_y \tau = 1, \\ &= \mathbf{T}_1 + \tau_1(..) = \mathbf{T}_2 + \tau_2(..) = \dots,\end{aligned}$$

we find $$D = \Sigma I_0\{\mathbf{T}_i, \tau_i\} = 2\Sigma I_0\{\tau_i, \tau_j\}, \dots \quad (i \neq j),$$

or, in the rather elliptical form often adopted, D is twice the number of O-intersections of F by itself.

As regards the determination of tangents from a point, the O-intersections of F with the polar of the point, D in number, are nugatory. Hence to find the number of O-tangents to a component Q from an arbitrary point on the O-tangent—that is, to find l, the class of the O-singularity on the component—if $y = 0$ is this tangent, and O_1 the arbitrary point, we have

$$l = I_0\{Q, \Delta_x Q\} - D.$$

As before

$$\begin{aligned}I_0\{Q, \Delta_x Q\} &= I_0\{\mathbf{T}, \Delta_x \mathbf{T}\} \\ &= I_0\{\mathbf{T}, \mathbf{T}_1 \Delta_x \tau_1 + \mathbf{T}_2 \Delta_x \tau_2 + \dots\} \\ &= \Sigma I_0\{\tau_1, \tau_2\tau_3 \dots \tau_k \Delta_x \tau_1\} \\ &= 2\Sigma I_0\{\tau_i, \tau_j\} + \Sigma I_0\{\tau_i, \Delta_x \tau_i\} \\ &= D + \Sigma I_0\{\tau_i, \Delta_x \tau_i\},\end{aligned}$$

so

$$l = \Sigma I_0\{\tau_i, \Delta_x \tau_i\}.$$

If now any

$$\begin{aligned}\tau &= y + A_\lambda x^\lambda + A_{\lambda'} x^{\lambda'} + \dots, \quad 1 < \lambda < \lambda' < \dots, \\ \Delta_x \tau &= \lambda A_\lambda x^{\lambda-1} + \lambda' A_{\lambda'} x^{\lambda'-1} + \dots \\ &= x^{\lambda-1}(\lambda A_\lambda + \dots),\end{aligned}$$

Digitized by Google

hence $$I_0\{\tau, \Delta_x\tau\} = I_0\{\tau, x^{\lambda-1}\} = \lambda - 1,$$

where $$\lambda = I_0\{\tau, y\},$$

and the summation over the k elements $\tau_1, \tau_2, \ldots \tau_k$ gives the result

$$l = \Sigma I_0\{\tau_i, y\} - k,$$

or $$k + l = \Sigma I_0\{\tau_i, y\} = I_0\{Q, y\},$$

that is, as stated earlier, the component and its tangent have $k+l$ O-intersections.

19. *Relation of F to the Hessian.* It is customary to complete the proof of Plücker's equations by means of the theorem of the invariance of the genus, p. This is doubtless the appropriate procedure when the discussion of $F(x, y) = 0$ is looked upon as a mere incident in the theory of algebraic functions; but it makes acquaintance with the individual curve dependent on knowledge of systems of curves, and while it is doubtless true that in order to know a part we must first know the whole, it is equally true that adequate comprehension of the whole is materially contributed to by detailed examination of the parts.

For a given curve F the determination of the elementary line-singularity, the inflexional tangent, is found to depend on the relation of F to its Hessian, an immediate consequence of the fact that the equation

$$\begin{vmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{vmatrix} = 0$$

includes the condition for an inflexion, though in a form not free from extraneous conditions. If this direct procedure is really in harmony with the spirit of the problem, it should be possible to apply it to the case of higher singularities. The result to be obtained may be divined from the relation (in which the summation is over the whole curve)

$$3\Sigma D - \Sigma\kappa + \Sigma\iota = 3m(m-2),$$

deduced from the direct expression of the theorem that the genus of $F(x, y)$ is equal to the genus of the reciprocal form. This relation states that if of the total number of intersections of F, H we set apart $\Sigma\iota$ as the number of *effective* intersections we are left with $\Sigma(3D - \kappa)$ *nugatory* intersections.

This suggests that for any singularity

$$I_0\{F, H\} = 3D - \kappa + \iota,$$

Digitized by Google

a relation easily proved for the elementary singularities and for others of fairly simple type, e.g. with either k or $l = 1$. The problem is to prove for higher singularities that N, the number of nugatory intersections, $= 3D - \kappa$. H. J. St. Smith pointed out the difficulty blocking the way to a direct solution of the question, the practical impossibility of finding the O-expansions for the Hessian. Halphen however at about the same date attacked the problem directly, using expansions for F, but for H the equation

$$H = 2F_1F_2F_{12} - F_1^2F_{22} - F_2^2F_{11} = 0,$$

the non-homogeneous equivalent of the determinant form of H in virtue of $z = 1$, $F = 0$. He relies on small quantities, considering orders of evanescence, segments of different orders, points in the vicinity of O. For the intersection of H with one expansion, τ, he says* that H reduces to $\dfrac{\partial^2\tau}{\partial x^2} F_y^3(x, y)_{\tau=0}$, in which $\dfrac{\partial^2\tau}{\partial x^2}$ is of order $\lambda - 2$ ($\epsilon - 1$ in Halphen), and therefore infinitely small of order $\lambda - 2$, finite, or infinitely great of order $2 - \lambda$ according as $\lambda >= < 2$. Hence the number of intersections on this branch is

$$3I_0\{\tau, F_y\} + \lambda - 2,$$

and therefore, assuming that the summation is not prohibited by the possible negative value of $\lambda - 2$ which makes certain quantities large instead of small, he finds

$$\begin{aligned} I_0\{F, H\} &= 3\Sigma I_0\{\tau, F_y\} + \Sigma(\lambda - 2) \\ &= 3I_0\{F, \Delta_y\} + l - k. \end{aligned}$$

This seems to be the only frontal attack that has been made on the general problem of determining the number of O-intersections of F and H. By means of the resolved form this line of argument can be followed without reliance on small quantities; a modification then suggests itself which obviates the apparent necessity—to which Halphen yielded—of a rather dubious if not actually unjustifiable extension of a convention.

20. It has been shown that the resolved form may be used in forming the Hessian. The formula for the Hessian of a product, viz.

$$H(F) = H(f_1f_2) = m_1f_2^3h_1 + m_2f_1^3h_2 + \Theta f_1f_2,$$

shows

$$\{F, H\} = \{f_1, f_2^3h_1\} + \{f_2, f_1^3h_2\},$$

so

$$I_0\{F, H\} = 6I_0\{f_1, f_2\} + I_0\{f_1, h_1\} + I_0\{f_2, h_2\}.$$

* It is not quite clear how Halphen himself arrived at this, but the context suggests that it was by means of small quantities.

Digitized by Google

If now the k factors $\tau_1, \tau_2, \ldots \tau_k$ in $\mathbf{T}$ are associated in any way, so that $\mathbf{T} = f_1 f_2 f_3 \ldots$, this gives

$$I_0\{\mathbf{T}, H\mathbf{T}\} = 6 \sum_{i \neq j} I_0\{f_i, f_j\} + \Sigma I_0\{f_i, h_i\}$$
$$= 3\bar{D} + \Sigma I_0\{f_i, h_i\},$$

where $\bar{D}$ is the contribution to the discriminantal index arising from the factors $f_1, f_2, f_3, \ldots$ (the components) in pairs, and $I_0\{f_i, h_i\}$ is to be determined for each component separately.

For a linear component

$$\tau = y + A_1 x + A_2 x^2 + \ldots + A_\sigma x^\sigma,$$

since $$\tau_y = 1, \quad \tau_{xy} = 0, \quad \tau_{yy} = 0,$$

the Hessian, given by the non-homogeneous form

$$2\tau_x \tau_y \tau_{xy} - \tau_x^2 \tau_{yy} - \tau_y^2 \tau_{xx} + \tau \mathbf{F} = 0,$$

reduces to $$\tau_y^2 \tau_{xx} + \tau \mathbf{F} = 0.$$

Hence $$I_0\{\tau, h\} = I_0\{\tau, \tau_{xx}\}.$$

If λ is the lowest exponent in τ after the terms $y + A_1 x$,

$$\lambda \geqslant 2; \ \lambda = l + k = l + 1.$$

Then $$\tau_{xx} = \lambda(\lambda - 1) A_\lambda x^{\lambda - 2} + \ldots,$$

which contains $x^{\lambda-2}$ as a factor, and

$$I_0\{\tau, h\} = I_0\{\tau, x^{\lambda-2}\} = \lambda - 2 = l - 1,$$

the inflexional index of τ. So the inflexional index of a linear component is actually the number of intersections with the Hessian.

For a superlinear component we have

$$Q = \tau_1 \tau_2 \ldots \tau_k,$$

where $$\tau_\lambda = y + A_1 x + \sum_\lambda^\sigma A_\mu x^\mu,$$

is one of a set of k conjugate elements.

Since $y + A_1 x = 0$ is the tangent to every τ in Q,

$$I_0\{Q, y + A_1 x\} = \Sigma I_0\{\tau, y + A_1 x\} = \sum_k \lambda = k\lambda.$$

But $$I_0\{Q, y + A_1 x\} = k + l,$$

hence $$k\lambda = k + l, \quad \lambda = 1 + \frac{l}{k}.$$

The formula for $I_0\{\mathbf{T}, H\mathbf{T}\}$ applied to Q expresses $I_0\{Q, H\}$ as $3D + \Sigma I_0\{\tau_i, h_i\}$, where h is the Hessian of τ. This is equivalent to the formula used by Halphen. Exactly as for the linear component τ, we obtain the purely formal result

$$I_0\{\tau, h\} = I_0\{\tau, \tau_{xx}\} = I_0\{\tau, x^{\lambda-2}\} = \lambda - 2.$$

Digitized by Google

If $\lambda \geqslant 2$, this gives a positive value, integral or fractional, for $I_0\{\tau, h\}$; hence

$$\begin{aligned} I_0\{Q, H\} &= 3D + \sum_k (\lambda - 2) \\ &= 3D + k(\lambda - 2) \\ &= 3D + l - k = 3D + \iota - \kappa, \end{aligned}$$

so $$N = 3D - \kappa.$$

If however $\lambda < 2$, the first term in τ_{xx} has a negative exponent. If we accept this formal result, without regard to the geometrical interpretation, we find a negative value for $I_0\{\tau, h\}$; the contribution $\Sigma I_0\{\tau, h\}$ to the total number of intersections presents itself not as an addition to $3D$ but as a correction to be applied. This suggests that the special arrangement adopted, which imposes on the desired number a purely arbitrary division into $3D$ and another part, though algebraically correct is actually inappropriate; it ignores some vital fact. Probably as a mere matter of algebraic arrangement the negative value is admissible; Halphen accepts this convention with its consequences, and modifies the $3D$ by the subtraction then indicated. But, just as in the theory of the simple intersections of curves, there arises the question whether we can properly, even as a mere numerical step, take away a number of intersections; we must first be assured that there is an available set of the required numerical value. Such a set presents itself here if in the first place we determine the intersections of Q with xH, for which we shall find, with the same arrangement of the algebraic work,

$$I_0\{Q, xH\} = 3D + l.$$

21. To show this, write $Q = \tau_1 T_1$, $T_1 = \tau_2 T_2$, etc.; then

$$H(Q) = M'\tau_1^3 H_1 + M''T_1^3 h_1 + \Theta Q,$$

hence $$xH = M'\tau_1^3 xH_1 + M''T_1^3 xh_1 + \Theta xQ,$$

so
$$\begin{aligned} I_0\{Q, xH\} &= I_0\{T_1, xH\} + I_0\{\tau_1, xH\} \\ &= I_0\{T_1, \tau_1^3 xH_1\} + I_0\{\tau_1, T_1^3 xh_1\} \\ &= 6I_0\{\tau_1, T_1\} + I_0\{T_1, xH_1\} + I_0\{\tau_1, xh_1\}. \end{aligned}$$

But $$T_1 = \tau_2 T_2,$$

hence $$I_0\{T_1, xH_1\} = 6I_0\{\tau_2, T_2\} + I_0\{T_2, xH_2\} + I_0\{\tau_2, xh_2\},$$

and so on.

Thus
$$\begin{aligned} I_0\{Q, xH\} &= 6\Sigma I_0\{\tau_i, T_i\} + \Sigma I_0\{\tau_i, xh_i\} \\ &= 3D + \Sigma I_0\{\tau_i, xh_i\}. \end{aligned}$$

Digitized by Google

In this $\tau = y + A_1 x + A_\lambda x^\lambda + \ldots + A_\sigma x^\sigma,$

so $xh = \lambda(\lambda - 1) A_\lambda x^{\lambda-1} + \ldots + \sigma(\sigma - 1) A_\sigma x^{\sigma-1},$

which contains $x^{\lambda-1}$ as a factor. Since $\lambda - 1$ is positive, this gives

$$I_0\{\tau, xh\} = I_0\{\tau, x^{\lambda-1}\} = \lambda - 1,$$

and $$\Sigma I_0\{\tau_i, xh_i\} = \Sigma(\lambda - 1) = l;$$

hence $$I_0\{Q, xH\} = 3D + l.$$

Now $$\{Q, xH\}_0 = \{Q, H\}_0 + \{Q, x\}_0;$$

in this $\{Q, x\}_0$ is a complete set, in number k, which may be removed. Consequently

$$\begin{aligned} I_0\{Q, H\} &= 3D + l \text{ diminished by } k \\ &= 3D + 1 + \iota \text{ diminished by } 1 + \kappa \\ &= (3D - \kappa) + \iota, \end{aligned}$$

so $$N = 3D - \kappa.$$

This shows that for the whole singularity at O

$I_0\{F, H\} = 3\bar{D}$ (due to combinations of the components in pairs)

$+ \Sigma(3D - \kappa) + \Sigma\iota$ (due to the separate components)

$= 3D - \kappa + \iota$, where D, κ, ι are the indices at O.

Hence for the whole curve

$$I\{F, H\} = 3\Sigma D - \Sigma\kappa + \Sigma\iota = 6\nu + 8\kappa + \iota,$$

so $$\iota = 3m(m - 2) - 6\nu - 8\kappa.$$

This completes the direct proof of Plücker's equations by finite processes and the use of finite quantities only. Throughout the work all the coefficients of F that affect the result have been retained; the apparent rejection of any terms is justified not on the ground that the aggregate of these terms is negligible, but for the more satisfactory reason that they simply are not present.

22. The terminated expansions here called elements were introduced by Brill*. Any such expansion defines parametrically a rational curve of a certain type,

$$x = t^k, \quad y = \sum_\lambda^\sigma A_\mu t^\mu,$$

with a singularity at O. Any curve F with a complicated singularity at O can be approximated to in the vicinity of O by a number of these components, "rational-ganze Curven"; and from these by small

* "Ueber Singularitäten ebener algebraischer Curven u.s.w. 1879," *Math. Annalen*, xvi. pp. 348–408; *Vorles. ü. ebene algebraische Kurven u.s.w.* 1925, pp. 263–272.

Digitized by Google

variation in the coefficients is obtained for F a penultimate form which has only elementary singularities, all separate. If in the vicinity of O this penultimate form has κ cusps, ι inflexions, ν nodes, τ pairs of contacts of bitangents, then the original singularity that F has at O is said to contain κ cusps etc. These are Brill's definitions of κ etc. as involved in a composite singularity. These numbers are to be found for a component from the degrees of the parametric equations for such singularities, from which $\kappa = k-1$, $\iota = \lambda - k - 1$. But as to the numbers ν, τ, Brill gives the proof only for a single component with k prime to l, the one case in which an *à priori* result can be obtained without detailed examination of the expansions, since the number of latent double elements, point or line, is $\frac{1}{2}(k-1)(l-1)$ in this case only. For the determination of ν, τ in the more general case for one component, and apparently also for these numbers as arising from the combination of components, he refers to the investigations of Smith and Halphen on the critical or characteristic exponents in the expansions. Inasmuch as the κ, ι, ν, τ as defined by Brill are thus determined, he finds no occasion for considering directly the relation of F and the Hessian at a higher singularity.

This use of the terminated expansions is entirely different from the use made of them above. The whole of Brill's interesting investigation, starting with the definition of the contents of a singularity, is frankly topological; both procedure and proof are based on the recognition of small quantities.

Digitized by Google

The (2, 1) *Correspondence.* By R. VAIDYANATHASWAMY. (Communicated by Mr H. W. TURNBULL.)

[*Received* 21 April, *read* 3 May, 1926.]

INTRODUCTION.

This communication is a sequel to a former paper on "The General (m, n) Correspondence" read before this Society in March, 1926*. It contained the general exposition (§ 1) and the theory of complete and closed sets (§ 2). For convenience these reference numbers will be here adopted, so that the present work, which opens with § 3, will be understood as a direct application of the previous sections.

The main object is to establish results which have a geometrical application in the geometry on the circle. In two elaborate memoirs† Pittarelli has studied the (2, 1) correspondence, where he develops the notions of what have here been called *apolar pair*, *auto-pair*, and *apolar cubic*. He proves the fundamental result that the *auto-pair* is the Hessian pair of the apolar cubic, but not the other fundamental result that the cubic of fixed points is apolar to the cubic-covariant of the apolar cubic [see § 3 (14)]. Beyond this, there is little in common, either in method or result, between the present work and the work of Pittarelli.

The properties of the *orthocentric pencil* lead to a generalized algebraic statement, explained in § 4 as the two-quadratic representation of the pencil of cubics.

In § 5 the binary geometry of the circle is developed. A fundamental link with what precedes is the covariant specification of the orthocentre of an inscribed triangle of the circle. For the triangle defines a certain (2, 1) correspondence on the circle, whose *apolar pair* in turn corresponds with the orthocentre.

CONTENTS.

§ 3. THE (2, 1) CORRESPONDENCE.
(10) Quadratic Covariants of the (2, 1) Form.
(11) Relation between the Auto-pair and the Apolar Pair for given fixed points.
(12) Apolar and Auto-pairs of special forms.
(13) Indeterminate cases of Apolar Pair.
(14) The Apolar Cubic. Similar Forms.
(15) Automorphic transformations.
(16) Canonical resolution.

* *Proc. Camb. Phil. Soc.*, vol. XXIII, pp. 109–119.

† *Atti della R. Accad. Lincei*, vol. III (1885–6). I must thank Mr F. P. White (St John's College) for having drawn my attention to these memoirs.

Digitized by Google

§ 4. The Two-Quadratic Representation of the Pencil of Cubics.
(17) Derivation of the representation.
(18) Representation of Apolar Pencils.
(19) Representation of the Null Pencil.
(20) Absolute Parameters of a Pencil.

§ 5. The Binary Geometry of the Circle.
(21) Special (2, 1) Correspondences on the Circle.
(22) The Pedal Angle.
(23) Polar correspondence related to the auto-correspondence.
(24) Pedo-parallel and Orthocentric Pencils.
(25) Orthocentric Loci.
(26) Other systems of Inscribed Triangles. The Antipedal Angle.

§ 3. The (2, 1) Correspondence.

(10) *Quadratic Covariants of the* (2, 1) *Form.*

The (2, 1) form and its quadratic covariants have a bearing of a fundamental character on the binary geometry of the circle, and their study is indispensable as a preliminary to that of the *pedal correspondence.*

Let the (2, 1) form $F(x, y)$ be written symbolically

$$F(x, y) = (ax)^2 (by) = (a'x)^2 (b'y) = (a''x)^2 (b''y).$$

The quadratic in x which is apolar to F for all values of y is evidently:

$$3{\cdot}1 \qquad O(x) = (aa')(bb')(ax)(a'x).$$

We call $O(x)$ the apolar quadratic of F. Since $O(x)$ is also the first rank covariant, or the (1, 1) transvectant of F, it follows* that:

3·2 *The apolar quadratic of F vanishes identically when and only when F is of rank one.*

By iterating the correspondence F, we can see that there exists only one pair of points ij, each of which corresponds to the other in F. Thus, if $x = p$ correspond in F to $y = q$, and $x = q$ to $y = r$, then the iterated correspondence of F, being the (4, 1) correspondence (p, r), has *five* fixed points, among which the three fixed points of F would be included. The two other fixed points ij are thus fixed points of the iterated, but not of the original correspondence. This can only mean that each of ij corresponds to the other in F, or ij is a *semi-complete* pair of F (§ 2 (6)). We shall call (ij) the *auto-pair* of F, and a constant multiple of $(ix)(jx)$ the *auto-quadratic* of F.

This iteration-process for determining ij amounts algebraically to the elimination of y between $(ax)^2 (by)$ and $(ay)^2 (bx)$, and the removal of the factor $f(x) = (ax)^2 (bx)$ from the result. The work

* Cf. § 1 (1). Also *Proc. London Math. Soc.* 2, 24 (1925), p. 83.

Digitized by Google

for the symbolic determination of the auto-quadratic would therefore run as follows:

$$(ax)^2 (by) = 0; \qquad \therefore \; y = b\,(ax)^2.$$

$$\begin{aligned}\therefore\ (a''y)^2 (b''x) &= (a''b)\,(ax)^2\,(a''b')\,(a'x)^2\,(b''x)\\ &= (ax)^2\,(a'x)^2\,(a''b')\,[(a''x)\,(b''b) + (a''b'')\,(bx)]\\ &= f(x)\,(a'x)^2\,(a''b')\,(a''b'')\\ &\quad + \tfrac{1}{2}\,(a'x)^2\,(ax)\,(a''x)\,(b''b)\,\{(a''b')\,(ax) - (a''x)\,(ab')\}\\ &= f(x)\,\{(ax)^2\,(a'b)\,(a'b') + \tfrac{1}{2}\,O(x)\} \quad (3{\cdot}1).\end{aligned}$$

Hence the auto-quadratic $S(x)$ may be taken to be:

3·3 $$S(x) = (ax)^2\,(a'b)\,(a'b') + \tfrac{1}{2}O(x).$$

The first term on the right of this is obtained by substituting $-a'(a'b')$ for y in $(ax)^2(by)$. Since $O(x)$ is apolar to $(ax)^2(by)$ for all values of y, it follows that $O(x)$ is apolar to $(ax)^2(a'b)(a'b')$. Hence from 3·3 we have:

3·4 *The apolar invariant* $(S(x), O(x))^2$ *is equal to a half of the discriminant of* $O(x)$.

Now, if $(ax)^2(a'b)(a'b') = 0$ identically, this would imply either that there is a value of y, namely, $y = -a'(a'b')$, which makes $(ax)^2(by)$ vanish identically, or that $(a'x)(a'b') = L(x)$ vanishes identically. If we suppose that F is not of rank one, then there can be no value of y which makes F vanish identically, so that the first alternative becomes impossible. Hence $L(x) \equiv 0$. But $L(x)$ is clearly the second form in the ladder of $F = (ax)^2(by)$ (1·12, 1·13); hence its identical vanishing implies that F is a polar form. Thus:

3·5 *If* $O(x) \neq 0$ *and* $(ax)^2(a'b)(a'b') \equiv 0$, F *must be a polar form.*

If $(ax)^2(bx) = f(x) = (\alpha x)(\beta x)(\gamma x)$, and $F(x, y)$ is put in its canonical form 1·2, namely,

$$F(x, y) = \mu_1(\alpha y)(\beta x)(\gamma x) + \mu_2(\beta y)(\gamma x)(\alpha x) + \mu_3(\gamma y)(\alpha x)(\beta x),$$

it is easy to show that the apolar and auto-quadratics are (except for constant multiples) given by:

3·6 $$\begin{cases} O(x) = \Sigma\,\mu_2\mu_3\,(\beta\gamma)^2\,(\alpha x)^2, \\ S(x) = \Sigma\,\mu_1\,(-\mu_1 + \mu_2 + \mu_3)\,(\gamma\alpha)\,(\alpha\beta)\,(\beta x)\,(\gamma x). \end{cases}$$

If the auto-pair ij or the apolar pair lm (namely, the pair of roots of the apolar quadratic) of F is given, it is easy to see that F could be written in the respective shapes:

3·7 $$F(x, y) = (ix)\,(iy)\,(Ax) + (jx)\,(jy)\,(Bx).$$

3·8 $$F(x, y) = (lx)^2\,(m'y) + (mx)^2\,(l'y).$$

Digitized by Google

Besides the apolar quadratic and auto-quadratic, two other quadratic covariants of F, of secondary importance, may be mentioned. These are first, the quadratic $O'(y)$ whose roots $l'm'$* are the values of y which correspond in F to its apolar pair, and secondly, the residual quadratic $R(x)$ whose roots A, B are the values of x other than ij which correspond in F to its auto-pair ij. $O'(y)$, being clearly the discriminant of F in respect of x, is given by:

3·9 $$O'(y) = (aa')^2 (by)(b'y).$$

The residual pair is composed of the harmonic conjugates of ij in respect of the apolar pair; hence

$$R(x) = 4O(x)\, D_{22} - 2S(x)\, D_{12},$$

where $$D_{22} = (S, S)^2, \qquad D_{12} = (O, S)^2.$$

Hence by 3·4:

3·10 $$R(x) = 4O(x)\, D_{22} - S(x)\, D_{11}.$$

(11) *Relation between Auto-pair and Apolar Pair for given fixed points.*

The knowledge of either the apolar pair or the auto-pair amounts to two linear conditions. Hence the (2, 1) correspondence F is uniquely determined when its apolar pair or its auto-pair is known in addition to its fixed points. Therefore for a given triad of fixed points $\alpha\beta\gamma$, the apolar pair and auto-pair determine each other uniquely.

Geometrically, let F be considered as a (2, 1) correspondence between the points of a fundamental conic C, the fixed points being $\alpha\beta\gamma$. The apolar and auto-pairs are then represented by two points O and S of the plane. It follows then, *that S is a certain Cremona (quartic) transformation Γ of O.* The transformations Γ, Γ^{-1} have some important properties which we may briefly mention here.

To obtain the equations to Γ, Γ^{-1}, it is best to use a binary coordinate system with $\alpha\beta\gamma$ as the triangle of reference. Any point X is the meet of tangents at two points of C, the binary parameters of which are determined by a quadratic of the form:

$$X_1 (\beta\gamma)^2 (\alpha x)^2 + X_2 (\gamma\alpha)^2 (\beta x)^2 + X_3 (\alpha\beta)^2 (\gamma x)^2.$$

We take $X_1 X_2 X_3$ as the coordinates of X. Similarly if the intersections of C with a line L are determined parametrically by a quadratic of the form

$$L_1 (\gamma\alpha)(\alpha\beta)(\beta x)(\gamma x) + L_2 (\alpha\beta)(\beta\gamma)(\gamma x)(\alpha x) + L_3 (\beta\gamma)(\gamma\alpha)(\alpha x)(\beta x),$$

we take $L_1 L_2 L_3$ to be the coordinates of L. The point-coordinates and line-coordinates thus chosen are mutually compatible. For,

* Called by Pittarelli (*loc. cit.*) "punti di diramazione." He gives a few of their properties.

Digitized by Google

the incidence condition of the point X and the line L is, from our definition, the apolarity condition of the corresponding quadratics, which is seen to be $L_1X_1 + L_2X_2 + L_3X_3 = 0$. The equation to the conic C in point-coordinates, being the condition that the quadratic

$$X_1(\beta\gamma)^2(\alpha x)^2 + X_2(\gamma\alpha)^2(\beta x)^2 + X_3(\alpha\beta)^2(\gamma x)^2$$

be a perfect square, is

$$2X_2X_3 + 2X_3X_1 + 2X_1X_2 = 0.$$

If we suppose the correspondence F expressed in the shape $\mu_1(\alpha y)(\beta x)(\gamma x) + \ldots$, then the apolar and auto-quadratics are given by 3·6. Hence the coordinates of the points O and S are:

$$O : (\mu_2\mu_3, \mu_3\mu_1, \mu_1\mu_2);$$

$$S : \{(\mu_1^2 - \overline{\mu_2 - \mu_3}^2), \quad (\mu_2^2 - \overline{\mu_3 - \mu_1}^2), \quad (\mu_3^2 - \overline{\mu_1 - \mu_2}^2)\}.$$

Therefore the equations to Γ are:

$$X_1' = X_2^2X_3^2 - X_1^2(X_2 - X_3)^2,$$
$$X_2' = X_3^2X_1^2 - X_2^2(X_3 - X_1)^2,$$
$$X_3' = X_1^2X_2^2 - X_3^2(X_1 - X_2)^2.$$

From these equations it may be shown that:

3·11 *A fundamental point of the first order, and one of the second order, of* Γ, *coalesce at each of* $\alpha\beta\gamma$, *so that these points correspond in* Γ *to line-pairs; the points* $\alpha\beta\gamma$ *are also fundamental points of the second order of* Γ^{-1}, *while the poles of* $\beta\gamma$, $\gamma\alpha$, $\alpha\beta$ *in respect of* C *are fundamental points of the first order of* Γ^{-1}; *the conic* C *is a curve of fixed points for* Γ.

These statements also become evident on decomposing Γ into simpler transformations. Thus it may be verified that Γ is the result of performing successively the following operations: (1) the operation **P** of taking the polar line in respect of the triangle $\alpha\beta\gamma$, (2) the operation **C** of taking the pole in respect of C, and (3) a certain quadratic involutory Cremona transformation I (Steinerian or generalised isogonal transformation) with the fundamental points $\alpha\beta\gamma$.

A triangle $\alpha\beta\gamma$ inscribed in a fundamental conic C, determines two noteworthy transformations. The first is the perspectivity V which carries $\alpha\beta\gamma$ respectively into the vertices of its polar triangle in regard to C. The centre and axis of the perspectivity V are pole and polar in regard to C, and correspond, as is well known, to the Hessian pair of $\alpha\beta\gamma$. The second transformation is the Cremona involution already called I. Among the ∞^1 quadratic involutions which have the fundamental points $\alpha\beta\gamma$ (these points corresponding to the opposite sides), the transformation I is distinguished by the fact that it has the centre of the perspectivity V for a fixed point

Digitized by Google

(the three other fixed points being the vertices of the polar triangle of $\alpha\beta\gamma$). Between these two transformations, and the transformations **P**, **C**, there obtains, as may be easily verified, the relation:

$$VI = \mathbf{CP}.$$

This relation replaces our decomposition of Γ by a more symmetrical decomposition*, namely

$$\Gamma = I\mathbf{CP} = IVI; \quad \Gamma^{-1} = IV^{-1}I.$$

(12) *Apolar and Auto-pairs of Special Forms.*

If F is of rank one, say $F = (ix)(jx)(\lambda y)$, the apolar quadratic $O(x)$ vanishes identically (3·2), and therefore the apolar pair is indefinite, and may be any pair apolar to ij. The auto-quadratic reduces now to its first term, which is seen to be a vanishing multiple of $(ix)(jx)$. The auto-pair is accordingly definite and is the pair ij. .....(3·12).

Consider next the case in which F has a binary linear factor (ax), say $F = (ax).(ax)(by)$. The apolar quadratic is then $\frac{1}{2}(ax)^2(aa')(bb')$ (3·1). By 3·3 the auto-quadratic is seen to be $(ax)(ax)(ab)(a'b')$. Thus if $(ax)(by)$ is a general form, the auto-pair may be any pair containing a; if, however, $(a'b') = 0$, that is, if $(ax)(by)$ is symmetric, the auto-quadratic vanishes, and the auto-pair may be any pair of the involution $(ax)(by) = 0$. These results are confirmed by 3·11. .....(3·13).

We shall frequently meet with problems in which it is required to know the nature of the form F, when certain relations hold between its apolar pair and auto-pair.

If the discriminant of the apolar quadratic is known to vanish, then, either $O(x)$ vanishes identically, so that F is of rank one (3·2), and we have the case 3·12, or the apolar quadratic is a perfect square $(ax)^2$. Since F is identically apolar to $O(x)$, it follows that in the latter case F develops the binary factor (ax), reducing to 3·13.

By 3·4, the apolarity of $S(x)$ and $O(x)$ amounts to the vanishing of the discriminant of $O(x)$. Hence from 3·12, 3·13, we have:

3·14 *The only case in which the auto-pair* ij $(i \neq j)$ *and the apolar pair of* F *are mutually apolar, without the latter being* necessarily *a repeated pair, is when* F *is of the shape* $(ix)(jx)(\lambda y)$.

One other case of importance is that in which the given auto-pair ij $(i \neq j)$ is known to be also the apolar pair. For this case the first term in the expression 3·3 for $S(x)$ must either vanish, or be a multiple of $O(x)$. Since this first term is apolar to $O(x)$, the latter alternative would reduce to the vanishing of the discriminant of $O(x)$, and therefore to 3·12, 3·13, in neither of which,

* I owe this symmetrical decomposition to a suggestion of Professor H. F. Baker.

Digitized by Google

however, the auto-pair and the apolar pair are identical. Hence, the first term must vanish, so that by 3·5 F is a polar form. It is clear that $F(x, x)$ has ij then for its Hessian pair. The same results may be reached by using 3·7, 3·8. .....(3·15).

(13) *Cases of Indeterminateness of Apolar Pair for given Auto-pair.*

A form F with a given auto-pair ij can be expressed in only one way in the shape 3·7, namely:

$$F(x, y) = (ix)(iy)(Ax) + (jx)(jy)(Bx).$$

The fixed cubic and the apolar quadratic of F are then:

$$3\cdot16 \qquad \begin{cases} f(x) = F(x, x) = (ix)^2(Ax) + (jx)^2(Bx), \\ O(x) = 2(ij)\{(ix)(Ax), (jx)(Bx)\}^1. \end{cases}$$

Hence it follows that $O(x)$ can vanish identically only in three cases: (1) $(Ax) = 0$ identically, (2) $(Bx) = 0$ identically, (3) $(ix)(Ax)$ and $(jx)(Bx)$ differ only by a constant factor, that is,

$$(iB) = (jA) = 0.$$

In the first case, $f(x)$ is $(jx)^2(Bx)$, and the apolar pair may be any pair apolar to jB (3·12). If, however, we specify a definite mode of approach to $f(x)$, the apolar pair becomes determinate. Thus, let

$$f(x) = (jx)^2(Bx) = \operatorname*{Lt}_{\delta \to 0} \{(jx)^2(Bx) + \delta\psi(x)\},$$

where $\psi(x) = (ix)^2(Cx) + (jx)^2(Dx)$ is an arbitrary cubic specifying the mode of approach. Then, from 3·16,

$$\begin{aligned} O(x) &= \{\delta(ix)(Cx), (jx)[(Bx) + \delta(Dx)]\}^1 \\ &= \delta\{(ix)(Cx), (jx)(Bx)\}^1 \text{ in the limit.} \end{aligned}$$

Removing the scalar factor δ, we see that the apolar pair has a definite limiting position. It is also clear that this limiting position depends only on the part $(ix)^2(Cx)$ of $\psi(x)$. Hence:

3·17 *When the fixed cubic $f(x)$ of a form with the auto-pair ij is of the shape $(jx)^2(Bx)$, the apolar pair may be any pair apolar to jB. If, however, we specify a direction of approach to $(jx)^2(Bx)$ by means of an arbitrary cubic $\psi(x)$, the apolar pair has a definite limiting position which is the same for all cubics ψ of the shape*

$$\psi(x) + (jx)^2(\lambda x).$$

A corresponding result holds for case (2). It is easy to obtain the following similar result for case (3):

3·18 *When the fixed cubic of a form with the auto-pair ij is of the shape $(ix)(jx)(Bx)$, the apolar pair is any pair apolar to ij; if, however, we specify a definite direction of approach to $(ix)(jx)(Bx)$*

Digitized by Google

by means of an arbitrary cubic $\psi(x)$, *the apolar pair has a definite limiting position which is the same for all cubics* ψ *of the shape* $\psi(x) + (ix)(jx)(\lambda x)$. (Cf. 3·12.)

The case in which the form F has ij for its auto-pair, and a perfect cube $(\alpha x)^3$ for its fixed cubic, is of special interest. Since this does not come under the cases (1), (2), (3) just discussed, it follows that the apolar pair of F is quite definite. To find it, we have the identity:

$$(ij)^3(\alpha x)^3 \equiv (j\alpha)^2(ix)^2\{(ij)(\alpha x) + 2(i\alpha)(jx)\} + (i\alpha)^2(jx)^2\{(ij)(\alpha x) - 2(j\alpha)(ix)\}.$$

Hence the form F is given by

$$F(x, y) = (ix)(iy)(Ax) + (jx)(jy)(Bx),$$

where

$$(Ax) = (j\alpha)^2\{(ij)(\alpha x) + 2(i\alpha)(jx)\},$$
$$(Bx) = (i\alpha)^2\{(ij)(\alpha x) - 2(j\alpha)(ix)\}.$$

The apolar quadratic of F is therefore:

3·19
$$O(x) = 2(ij)\{(ix)(Ax), (jx)(Bx)\}^1$$
$$= (ij)^2(i\alpha)^2(j\alpha)^2\{3(ij)^2(\alpha x)^2 + 4(i\alpha)(j\alpha)(ix)(jx)\}.$$

(14) *The Apolar Cubic. Similar Forms.*

Let the values of x which correspond to y in F be x_1x_2. Consider the system Γ of the ∞^1 triads yx_1x_2. If one point z of a triad of this type is known, the triad itself is determined in two ways, according as we consider z as an x or a y. Hence Γ is a quadratic system. Also, Γ is a continuous system. Since a continuous curve of the second order in space is necessarily a plane curve, it follows that all triads of Γ are apolar to a certain cubic $f''(x)$, which we call *the apolar cubic* of F.

Since $f''(x)$ is by definition identically apolar to

$$(xy)F = (xy)(ax)^2(by),$$

it follows that $f''(x)$ is the second rank covariant of the (3, 2) form $(xy)(ax)^2(by)$. This would effect the symbolic determination of the apolar cubic.

If ij is the auto-pair of F, and AB the residual pair (3·10), then each of the triads ijA, ijB is of the type yx_1x_2 and therefore apolar to $f''(x)$. Hence:

3·20 *The auto-pair of* F *is the Hessian pair of its apolar cubic* $f''(x)$.

If F is a polar form with the fixed points $\alpha\beta\gamma$, and if $\alpha''\beta''\gamma''$ are the cubicovariant points of $\alpha\beta\gamma$, then $y = \alpha$ corresponds in F to $x = \alpha, \alpha''$. Hence the apolar cubic of F is apolar to the three triads of the type $(\alpha\alpha\alpha'')$, and therefore represents the triad $\alpha''\beta''\gamma''$. Thus:

Digitized by Google

3·21 *The apolar cubic of a polar form F is the cubicovariant of its fixed cubic $f(x) = F(x, x)$.*

DEF. *Two forms F, F' will be said to be* SIMILAR, *if their apolar cubics differ only by a constant factor.*

It follows that all forms similar to F have the same auto-pair ij as F (3·20). It also follows from the definition, that $\lambda_1 F_1 + \lambda_2 F_2$ is similar to F, if F_1 and F_2 are similar to F. Hence the family of forms similar to F is a linear family, involving clearly three parameters. Hence there exists a polar form P similar to F. Also, if ij is the auto-pair of F, the apolar cubic of $(ix)(jx)(\lambda y)$ is indeterminate, and may be any cubic apolar to $(ix)(jx)$. Therefore the forms $(ix)(jx)(\lambda y)$ must be reckoned among the forms similar to F. We have therefore finally the result:

3·22 *If P is the polar form similar to F, the most general form similar to F is of the shape $\mu P(x, y) + (ix)(jx)(\lambda y)$, where ij is the auto-pair of F and μ, λ_1, λ_2 are arbitrary.*

Since F is similar to itself, it follows that F itself must be of this shape. Hence $f(x) = F(x, x) = \mu P(x, x) + (ix)(jx)(\lambda x)$. Let now $f''(x)$ be the apolar cubic of F, and $f'(x)$ the cubicovariant of $f''(x)$.

Since P is similar to F, its apolar cubic is a multiple of $f''(x)$; therefore since P is a polar form, it follows from 3·21 that $P(x, x)$ is a multiple of $f'(x)$. Also, by 3·20, ij is the Hessian pair of $f''(x)$, and therefore also of $f'(x)$, so that $(ix)(jx)(\lambda x)$ is apolar to $f'(x)$ for any λ. Hence:

3·23 *The cubic of fixed points $f(x) = F(x, x)$ is apolar to the cubicovariant of the apolar cubic of F.*

From these theorems we can construct the correspondence F, when its apolar triad $f''(x)$ and its apolar pair lm are given. Let x be any point, x' the harmonic conjugate of x in respect of lm, and $x_1 x_2$ the pair harmonic to xx' and ij, where ij is the Hessian pair of $f''(x)$, and therefore the auto-pair of F. Let y be the point whose first polar pair in respect of $f''(x)$ is $x_1 x_2$. Then y corresponds to x (and x') in F.

(15) *Automorphic Transformations of the* (2, 1) *Correspondence.*

In this and the next subsection, we consider correspondences between x, y, as *operations* which transform x into y; thus the correspondence $F(x, y) = 0$ would be indicated in the operational symbolism, by the equation:

$$y = Fx; \text{ or } x = F^{-1}y^*.$$

* I omit brackets after operational symbols F, F^{-1}, to avoid confusion with functional notation.

Digitized by Google

If F, F' be two correspondence-operations, the product FF' indicates as usual the result of performing F', F in succession.

The involution I whose fixed pair is the apolar pair lm of the (2, 1) correspondence F, will be called *the basic involution* of F. Since lm is by definition apolar to all pairs $F^{-1}y$, it follows that $Fx = FIx$. In other words:

3·24 *A* (2, 1) *correspondence* F *absorbs its basic involution* I*; in symbols* $F = FI$.

Two correspondences F, F' will be said to be *congruent*, if they have the same basic involution. Now a (2, 1) correspondence $F(x, y)$ may be considered as a (1, 1) correspondence between y and the pairs of the basic involution of F. Hence it is clear, that by following up F by a homography, we would obtain the most general correspondence congruent to F. Hence:

3·25 *Any two congruent correspondences* F, F' *are connected by a relation of the form* $F' = HF$, *where* H *is a homography.*

We shall now obtain the condition that the correspondence FH, where H is a homography, may be congruent to F. If I is the basic involution of F, then:

$$FH.H^{-1}IH = FIH$$
$$= FH \text{ (3·24)}.$$

Hence $H^{-1}IH$ is the basic involution of FH. Thus, for F and FH to be congruent, the involutions I, $H^{-1}IH$ must be the same. Now the fixed pair of I is the apolar pair lm of F, and the fixed pair of $H^{-1}IH$ is the pair $(H^{-1}l, H^{-1}m)$. Hence if $I = H^{-1}IH$, H is either a homography H_1 with the fixed points lm, or an involution I_1 containing the pair lm. The transformations H_1 and I_1 together form the discontinuous automorphic linear group of the point-pair lm, of which the continuous group $[H_1]$ is the maximum self-conjugate sub-group*. Thus:

3·26 *The homographies* H *which make* FH *congruent to* F, *are members of the automorphic linear group* $[H_1, I_1]$ *of the apolar pair of* F.

It should be noticed that $[H_1]$ contains a member, namely the basic involution I, which is permutable with every operation of $[H_1, I_1]$.

Now suppose that H has been so chosen that FH is congruent to F. Then by 3·25, there must be a relation of the form

3·27 $$F = H'FH,$$

* These are characteristic features of the automorphic linear group of a non-singular quadric in space of any odd number of dimensions, and are consequences of the existence of two algebraically distinct systems of generating regions on the quadric. In space of one dimension, the quadric is the point-pair, and the two points which constitute the pair answer to the two algebraically distinct systems of generating regions.

Digitized by Google

where H' is a homography. It is clear from this relation, that H' is a homography which transforms the pair $l'm'$ into itself, where $l' = Fl$, $m' = Fm$ are the correspondents of the apolar pair in F. Thus the homographies H' generate the automorphic linear group $[H'] = [H_1', I_1']$ of the point-pair $l'm'$. The groups $[H']$, $[H]$ are thus of the same type.

Now in the relation 3·27, H determines H' uniquely; also if H is of the type H_1, H' is also of the type H_1', and if H is of the type I_1, H' is of the type I_1'. Thus 3·27 is composed of two types of relations, $F = H_1'FH_1$, and $F = I_1'FI_1$, which we shall refer to as the *direct* and *skew* automorphic transformations of F respectively.

While H determines H' in 3·27, H' does not, however, determine H uniquely. The correspondence between H', H is such that when H' corresponds to H and $\bar{H}'$ to $\bar{H}$, then $H'\bar{H}'$ corresponds to $\bar{H}H$, and $\bar{H}'H'$ to $H\bar{H}$. Also, when H' is identity, then H is either identity or the basic involution I. Thus, it follows that the correspondence between H'^{-1} and H is a (1, 2) isomorphism. If H' corresponds to H, the other value of H which corresponds to H' is therefore $HI = IH$. We proceed to determine more precisely the (1, 2) correspondence (H', H), for the direct and skew transformations of F.

For the skew transformations $F = I_1'FI_1$, it is easy to see that when I_1' is a singular involution with the fixed pair $(l'l')$, then the two corresponding involutions I_1 both coalesce with the singular involution whose fixed pair is (ll). It is further evident on inspection, that when I_1 contains the auto-pair of F, I_1' also contains the auto-pair. Hence if the involutions I_1', I_1 are parametrically represented by η, ξ respectively, in such wise that the singular involutions correspond to the values 0, ∞ of the parameters, and the involutions I_1', I_1 which contain the auto-pair to the values $\eta = 1$, $\xi = 1$, then the correspondence between I_1', I_1 would clearly be expressed by the relation $\eta = \xi^2$.

Consider next the direct transformations $F = H_1'FH_1$. The operations of the Abelian group $[H_1]$ can be represented by an angular parameter θ, defined modulo π. Thus, if ν is the multiplier of H_1 at the fixed point l (say), we define θ to be $\frac{1}{2\sqrt{-1}} \log \nu$. If the angular parameters θ', θ of H_1', H_1 are reckoned in respect of corresponding fixed points l', l (say), then the (1, 2) correspondence (H_1', H_1) is determined by the following theorem:

3·28 *If $F = H_1'FH_1$ is any direct transformation of F, then the angular parameters θ', θ of H_1', H_1 are connected by the relation* $\theta' = -2\theta$.

For, consider a point $l + \delta z$, where δ is a small scalar, and z is an arbitrary fixed point. Since l is a fixed point of the basic

Digitized by Google

involution of F, it follows that $x = l + \delta z$ corresponds in F to a point which differs from l' by a small quantity of the second order. Thus

$$F\overline{l + \delta z} = l + \mu\delta^2 z,$$

where μ is independent of δ. Let now ν, ν' be the multipliers of H_1, H_1' at l, l'. Then:

$$\begin{aligned} H_1'FH_1\overline{l + \delta z} &= H_1'F\overline{l + \nu\delta z} \\ &= H_1'\overline{l' + \mu\nu^2\delta^2 z} \\ &= l' + \mu\nu'\nu^2\delta^2 z. \end{aligned}$$

Hence, since $F = H_1'FH$, $\mu\nu'\nu^2\delta^2 = \mu\delta^2$ or $\nu' = \frac{1}{\nu^2}$.

Thus $\theta' = \frac{1}{2\sqrt{-1}}\log\nu' = \frac{1}{2\sqrt{-1}}(-2\log\nu) = -2\theta.$

(16) *Canonical Resolution of the* (2, 1) *Correspondence.*

There are ∞^1 polar correspondences which are congruent to F; namely, those defined by the ∞^1 triads which have as their Hessian pair the apolar pair lm of F. Let P_1 be any one of these polar correspondences. Then by 3·25 there is a relation of the form

$$F = HP_1,$$

where H is a homography. Since $P_1 l = m$, $P_1 m = l$, it follows that $m' = Fm = Hl$, $l' = Fl = Hm$. Hence H can be expressed in only one way in the shape IH_1, where H_1 is a homography with the fixed points lm, and I is the involution determined by the pairs lm', $l'm$. Thus $F = IH_1P_1$; but $P = H_1P_1$ is only another of the ∞^1 polar correspondences congruent to F (since $H_1P_1 l = m$, $H_1P_1 m = l$). Hence:

3·29 *The correspondence F can be expressed uniquely in the shape IP, where P is a polar correspondence, and I an involution.*

We call I and P the *involutoric* and *polar components* of F, respectively. These components are related in a simple manner to the correspondence F' which has the same fixed points as F, and which has for its auto-pair the apolar pair of F.

Def. *F will be said to be the correlate of F', if F and F' have the same fixed points, and the apolar pair of F is the auto-pair of F'.*

With this definition we have the following theorem which specifies the two components of F:

3·30 *If F is the correlate of F', the involutoric component of F is the basic involution of F', and the polar component of F is the polar correspondence defined by the apolar cubic of F'.*

To prove this, let I be the basic involution of F', and P the polar correspondence defined by $f'''(x)$, the apolar cubic of F'.

Digitized by Google

Let a, b, c be the fixed points of F' and a', b', c' the *other* points which correspond in F' to $y = a, b, c$ respectively. From the definition of the apolar cubic it follows that each of the triads aaa', bbb', ccc' is apolar to $f''(x)$; therefore $Pa = a'$, $Pb = b'$, $Pc = c'$. But since $x = a$, a' correspond to the same value a of y in F', it follows that aa', bb', cc' are pairs of the basic involution I. so that $IPa = Ia' = a$. Thus the fixed points a, b, c of F' are also the fixed points of IP. Also, from the definition of P, the apolar pair of IP, which is the same as the apolar pair of P, is the Hessian pair of $f''(x)$, and therefore the auto-pair of F' (3·20). Hence IP is the correlate of F', as was to be proved.

The correlate F of the correspondence F' with the auto-pair ij may be constructed by a simple rule. Let x_1x_2 correspond in F to y, so that from the definition of correlate, x_1x_2 is a pair apolar to ij. Then if J be the involution whose fixed pair is x_1x_2, two of the fixed points of JF' are clearly ij. Let the third fixed point be y'. The correspondence between y' and the pairs x_1x_2, apolar to ij, is clearly (1, 1). Also, when x_1x_2 is one of the three pairs aa', bb', cc', $y' = y = a$, b, or c respectively. Hence $y' = y$ always; in other words:

3·31 *If J is a variable involution containing the auto-pair ij of F', the correlate of F' is simply the correspondence between the fixed pair of J and the fixed point other than ij of JF'.*

§ 4. The Two-Quadratic Representation of the Pencil of Cubics.

(17) *Derivation of the Representation.*

It was mentioned that the assigning of either the apolar pair or the auto-pair of a (2, 1) correspondence F, amounts to two linear conditions on the coefficients of F. Hence it follows that:

4·1 *The totality of* (2, 1) *forms $F(x, y)$ which have a given auto-pair $P_1(x)$, and a given apolar pair $P_2(x)$, belong to a pencil. Therefore also, the fixed cubics $F(x, x)$ of such forms belong to a pencil, which is completely determined by P_1, P_2.*

The pencil to which $F(x, x)$ belongs may therefore be denoted by the symbol $[P_1, P_2]$. Now there are ∞^4 pencils of cubics, and there are also ∞^4 pairs of point-pairs P_1, P_2. Hence any pencil of cubics can be represented in a finite number of ways in the form $[P_1, P_2]$. A representation of this kind we shall call "*a two-quadratic representation*" of the pencil.

4·2 *The quadratic $P_1(x)$ is a factor of the Jacobian of the pencil $[P_1, P_2]$.*

Digitized by Google

For, let $P_1(x) = (ix)(jx)$. The family of correspondences which have the auto-pair ij can be expressed in the shape 3·7, namely:

$$(ix)(iy)(Ax) + (jx)(jy)(Bx).$$

There is a unique member of this family, say F, which has P_2 for its apolar pair, and a fixed point at i. For i to be a fixed point of $F(x, y)$, B must be equal to ki, where k is numerical. If now $F(x, y)$ were of full rank two, its apolar quadratic would be a multiple of $k(ix)^2$, and therefore its apolar pair would not be that given by $P_2(x)$. Hence k must be zero, so that F reduces to $(ix)(iy)(Ax)$, where for P_2 to be the apolar pair, iA must be apolar to P_2. Thus $F(x, x) = (ix)^2(Ax)$, showing that i and similarly j are roots of the Jacobian of the pencil denoted by $[P_1, P_2]$. We have also obtained here a rule for constructing the pencil $[P_1, P_2]$, namely:

4·3 *Let* $P_1(x) = (ix)(jx)$, *and let the harmonic conjugates of* i, j *with respect to* P_2 *be* A, B *respectively. Then the pencil* $[P_1, P_2]$ *is simply the pencil determined by the two cubics*

$$(ix)^2(Ax), \quad (jx)^2(Bx).$$

Therefore, to obtain a two-quadratic representation of any pencil of cubics, we can take any quadratic factor of its Jacobian for P_1, and then construct P_2 by means of 4·3. Thus:

4·4 *There are exactly six two-quadratic representations of a pencil of cubics.*

These six representations fall into three *conjugate* pairs; namely, the representations $[P_1, P_2]$, $[P_1', P_2']$ are *conjugate*, if the Jacobian of the pencil is P_1P_1'.

(18) *Representation of Apolar Pencils**.

To see how the two-quadratic representations of two apolar pencils are mutually connected, we have first to specify the relation of the pencil $[P_1, P_2]$ to the quadratic apolar to P_1, P_2. For this purpose we begin by recalling some fundamental properties of a pencil Γ of cubics.

The pencil Γ possesses several quartic combinants, those which concern us here being: *the apolar quartic*, which is the unique quartic apolar to every member of Γ; *the Jacobian quartic*, the roots

* Two pencils of cubics Γ, Γ' are *apolar pencils*, if any member of either is apolar to any member of the other. Γ is a *null pencil*, if any two of its members are apolar.

If binary cubic forms are represented by points in three-space (as in sub-section 26 *infra*), the forms which are perfect cubes will correspond to a certain twisted cubic C. The null pencils will then correspond to the lines which belong to the tangent linear complex T of C; apolar pencils will correspond to polar lines of the complex T. It is well known that two polar lines of T are met by the same four tangents of C; the meaning of this is easily seen to be that apolar pencils have the same Jacobian.

Digitized by Google

of which are the repeated roots of members of Γ, and *the residual quartic*, the roots of which are the non-repeated roots of those members of Γ which possess repeated roots. Since Γ determines and is determined uniquely by its apolar quartic, it follows that all the quartic combinants must be covariants of the apolar quartic, and must, as such, belong to its syzygetic pencil*. The sextic covariant of this associated syzygetic pencil, which is, as is well known, the product of three mutually apolar quadratics, may be called *the sextic combinant* of Γ.

The three involutions determined by the three mutually apolar quadratic factors of the sextic combinant of Γ form with identity a group G of order four, which transforms every quartette of the associated syzygetic pencil into itself. Hence G must transform Γ into itself, and therefore every member of Γ into another member. In particular, G transforms a member with a repeated root into another member with a repeated root. Thus if $i_1 i_2 i_3 i_4$ be the Jacobian quartette, and $A_1 A_2 A_3 A_4$ the residual quartette, so that $(i_r x)^2 (A_r x)$ $(r = 1, 2, 3, 4)$ are the members of Γ with a repeated root, it follows that:

4·5 *Any involution of G effects the* SAME *permutation on the two quartettes $i_1 i_2 i_3 i_4$, $A_1 A_2 A_3 A_4$ which are the Jacobian and residual quartettes, written in the order in which they are associated in* Γ.

Now any two quartettes $i_1 i_2 i_3 i_4$, $A_1 A_2 A_3 A_4$, of a syzygetic pencil, when written down in arbitrary orders, may either be *isomorphic* or *antimorphic* in regard to the group G of the pencil; they would be isomorphic if any involution of G effects the *same* permutation on them both, antimorphic otherwise. The theorem 4·5 amounts, therefore, to the isomorphism of the Jacobian and residual quartettes, when in the orders in which they are associated in Γ. This isomorphism clearly implies that the quadratic apolar to $(i_r x)$ $(i_s x)$ and $(A_r x)$ $(A_s x)$ is one of the three mutually apolar quadratic factors of the sextic covariant, for $r, s = 1, 2, 3, 4$; $r \neq s$.

Now take any one of the six two-quadratic representations of Γ, say, the representation $[P_1, P_2]$, where $P_1 (x) = (i_1 x)(i_2 x)$. By 4·3 the quadratic P_2 is then the quadratic apolar to $(i_1 x)$ $(A_1 x)$ and $(i_2 x)$ $(A_2 x)$. The quadratic apolar to P_1 and P_2 is then seen to be simply the quadratic apolar to $(i_1 x)$ $(i_2 x)$ and $(A_1 x)$ $(A_2 x)$. Hence, by the theorem of isomorphism of order (4·5), it is one of the quadratic factors of the sextic combinant of Γ. We therefore have the result:

4·6 *The quadratic apolar to P_1 and P_2 is one of the three mutually apolar quadratic factors of the sextic combinant of the pencil $[P_1, P_2]$.*

* For properties of the syzygetic pencil of quartics, see Grace and Young, *Algebra of Invariants*, pp. 197–208. For the study of the syzygetic pencil in relation to pencils of cubics, Meyer's *Apolarität* may be consulted; Study, "On the irrational Covariants of certain Binary Forms," *Amer. Journ. of Math.* 17 (1895) will also be found useful.

Digitized by Google

It is well known that apolar pencils of cubics have the same Jacobian, and therefore the same sextic combinant. Therefore they could be simultaneously represented in the shapes $[P_1, P_2]$, $[P_1, P_3]$, P_1 being any factor of their common Jacobian. Now, there is only one quadratic factor of the sextic combinant, which is apolar to P_1. This quadratic factor is by 4·6 apolar to P_2 as well as to P_3. Thus there is a linear relation between P_1, P_2, P_3. Hence:

4·7 *Apolar pencils of cubics admit of simultaneous two-quadratic representations in the shapes* $[P_1, P_2]$, $[P_1, P_1 + kP_2]$.

(19) *Representation of the Null Pencil.*

The last theorem 4·7 points to the necessity of studying the family of pencils whose two-quadratic representations are of the shape $[P_1, P_1 + \mu P_2]$. Let us call the pencil $[P_1, P_1 + \mu P_2]$ the pencil $\Gamma\mu$. There are three fundamental μ-involutions which have to be considered; namely:

(i) the involution I_1 between μ_1, μ_2, such that the quadratics $P_1 + \mu_1 P_2$, $P_1 + \mu_2 P_2$ are apolar. The fixed elements of this involution correspond to the perfect squares of the pencil $P_1 + \mu P_2$, and are therefore given by

$$\mu^2 d_{22} + 2\mu d_{12} + d_{11} = 0,$$

where the d's are the usual invariants of P_1, P_2.

(ii) It is seen from 4·7 that the apolar pencil of a pencil $\Gamma\mu_1$ is a pencil $\Gamma\mu_2$. We call the involution between μ_1, μ_2 the involution I_2. The fixed elements of I_2 are the values of μ for which $\Gamma\mu$ is a null pencil; these values have yet to be determined.

(iii) We define the absolute invariant of two quadratics P_1, P_2 to be the number $\frac{d_{11}d_{22}}{d_{12}^2}$; thus the absolute invariant vanishes when either of the quadratics is a perfect square, becomes infinite when they are apolar, and takes the value unity when they have a common factor*. Since the absolute invariant is of the second degree in the coefficients of either quadratic, there must be two values of μ for which the absolute invariant of $P_1, P_1 + \mu P_2$ is a

* Let p' be one cross ratio of the point-pairs P_1, P_2, so that the other is $\frac{1}{p'}$. Then it is easily shown that the absolute invariant of (P_1, P_2) is $\frac{p-2}{p+2}$, where $p = p' + \frac{1}{p'}$. Again, if q' be a cross ratio of (P_1, P_2) and the pair of perfect squares in their pencil, and $q = q' + \frac{1}{q'}$, the absolute invariant may be shown to be $\frac{4}{q+2}$. If we represent quadratics as points in the plane of a fundamental circle of centre S and radius ρ, the points on the circle representing the perfect squares, then the absolute invariant of the quadratics represented by S and K may be shown to be $1 - \mu^2$, where $SK = \mu\rho$.

Digitized by Google

given number k. We define I_3 to be the involution of pairs of such values of μ. It is immediately verified that the two values of μ coalesce only when $k = \infty$ or 1. Now the absolute invariant of P_1, $P_1 + \mu P_2$ is equal to 1 only when P_1 and $P_1 + \mu P_2$ have a common factor, that is, only when $\mu = 0$; and the same absolute invariant is infinite only when $P_1 + \mu P_2$ is the quadratic P_1' of the pencil, which is apolar to P_1. Thus the fixed elements of I_3 correspond to the quadratics P_1, P_1'.

But since the quadratics P_1, P_1' are apolar, they correspond to a pair of the involution I_1; thus the fixed elements of I_3 belong to I_1. The two involutions I_1, I_3 are therefore symmetrically related, and are permutable with each other.

We next show that the fixed elements of I_3 belong also to I_2. For this we have to show that the pencils, whose two-quadratic representations are $[P_1, P_1]$, $[P_1, P_1']$ are apolar pencils. By 3·15 the (2, 1) correspondences whose auto-pair and apolar pair are identical with P_1 are the polar correspondences whose fixed triads are apolar to P_1. Hence the pencil $[P_1, P_1]$ is the pencil of triads apolar to P_1. By 3·14 the correspondences whose auto-pair is P_1, and whose apolar pair is apolar to P_1, are of the shape $P_1(x)(\lambda y)$, where λ is arbitrary. Hence the pencil $[P_1, P_1']$ is the singular pencil $P_1(x)(\lambda x)$. Thus the pencils $[P_1, P_1]$, $[P_1, P_1']$ are apolar pencils, so that the fixed elements of I_3 belong also to I_2. The involution I_3 may therefore be specified as that which contains the fixed pairs of I_1 and I_2.

To specify I_2 itself, we write every quadratic $P_1 + \mu P_2$ in the shape $P_1 + \lambda P_1'$; the perfect squares of the pencil $P_1 + \mu P_2$ are then of the shape $P_1 \pm tP_1'$, where t is given by $t^2 d' + d = 0$, d, d' being the respective discriminants of P_1, P_1'. Now write

$$P_1 + tP_1' = (ax)^2,$$

and consider the apolar pencil of the pencil $[P_1, (ax)^2]$. The (2, 1) form which has $(ax)^2$ for apolar pair, and $P_1(x)$ for its auto-pair must possess the binary factor (ax), as was seen in (12). Thus the members of the pencil $[P_1, (ax)^2]$ have the common factor (ax), so that its apolar pencil contains the member $(ax)^3$. We see then that the apolar pencil of $[P_1, (ax)^2]$ is that pencil of the family

$$[P_1, P_1 + \mu P_2] = \Gamma\mu,$$

which contains the member $(ax)^3$. Now the apolar pair of the (2, 1) form which has the auto-pair $P_1(x) = (ix)(jx)$, and the fixed cubic $(ax)^3$ has been determined in 3·19 to be

$$\begin{aligned} 3(ij)^2(ax)^2 + 4(ia)(ja)(ix)(jx) &= -6d(P_1 + tP_1') + 4dP_1 \\ &= -2d(P_1 + 3tP_1'). \end{aligned}$$

Thus the two pencils $[P_1, P_1 + tP_1']$, $[P_1, P_1 + 3tP_1']$ are apolar pencils; so also are the two pencils $[P_1, P_1 - tP_1']$, $[P_1, P_1 - 3tP_1']$.

Digitized by Google

The involution I_2 therefore contains the two pairs $\lambda = t,\ 3t$ and $\lambda = -t,\ -3t$; hence its fixed pair is given by $\lambda = \pm\sqrt{3}t$. This determines the null pencils of our family, namely:

4·8 *If the perfect squares of the pencil* $P_1 + \mu P_2$ *are* $P_1 \pm tP_1'$, *the null pencils of our family are the pencils* $[P_1,\ P_1 \pm \sqrt{3}tP_1']$.

Since $t^2d' + d = 0$, the absolute invariant of $P_1,\ P_1 \pm \sqrt{3}tP_1'$ is equal to $\dfrac{d\ (d + 3t^2d')}{d^2} = -2$. Hence:

4·9 *If* $[P_1, P_2]$ *is a null pencil, the absolute invariant of* P_1 *and* P_2 *is equal to* -2.

The results of this subsection give the fundamental properties of the two-quadratic representation. It is easy to derive from them any formulae or invariant expressions that may be required.

(20) *The Absolute Parameters of a Pencil.*

If the pencil Γ has the two-quadratic representation $[P_1, P_2]$, then it is clear that the absolute invariant of P_1, P_2 is an irrational absolute invariant of the pencil Γ; we shall call it *an absolute parameter* of Γ.

Suppose the Jacobian of Γ to be the product of the quadratics P_1, P_1', and let the corresponding conjugate representations of Γ be $[P_1, P_2]$, $[P_1', P_2']$. It is clear from (18), that two involutions of the group G associated with Γ transform P_1, P_2 simultaneously into P_1', P_2'. Thus the absolute invariants (P_1, P_2), (P_1', P_2') are equal, so that the six two-quadratic representations of Γ give rise only to three distinct absolute parameters.

It is clear that two pencils Γ, Γ' are linearly transformable into one another, if one absolute parameter of either is equal to one absolute parameter of the other. It follows from this, that the absolute invariant $t\left(= \dfrac{i^3}{6j^2}\right.$ in the usual notation for the invariants$\left.\right)$ of the apolar quartic of the pencil Γ must be a rational function of the absolute parameter s of Γ. We proceed to determine the exact form of this function.

If t is given, the pencil Γ is determined apart from an arbitrary linear transformation, and therefore its absolute parameter s is determined in three ways. Hence the relation between t and s is of the form

$$t = \frac{f(s)}{\phi\ (s)},$$

where $f\ (s)$, $\phi\ (s)$ are cubics. Now if $t = 0$, the apolar quartic of Γ is self-apolar and therefore Γ is a null pencil. But 4·9 shows that the absolute parameters of a null pencil are all equal to -2;

Digitized by Google

hence $f(s) = (s+2)^3$. Again, when $t = 1$, the apolar quartic of Γ has a squared factor $(\alpha x)^2$, and therefore the pencil Γ contains the perfect cube $(\alpha x)^3$. In this case it is easy to see that for four of the six representations $[P_1, P_2]$ of Γ, P_1 and P_2 have the common factor (αx). Thus when $t = 1$, two of the values of s are equal to 1. Lastly, when $t = \infty$, the apolar quartic has vanishing j and possesses an apolar quadratic $P_1(x)$. The pencil Γ is then the singular pencil $P_1(x)(\lambda x)$, and has therefore the representation $[P_1, P_1']$ (where P_1' is apolar to P_1), which counts four times among its six representations. Hence when $t = \infty$, two of the values of s are infinite. These conditions determine $\phi(s)$ as being equal to $27s$. We have therefore finally the relation:

$$4\cdot10 \qquad t = \frac{(s+2)^3}{27s},$$

between the absolute parameter of Γ and the absolute invariant of its apolar quartic*.

§ 5. The Binary Geometry of the Circle.

(21) *The Special* (2, 1) *Correspondences on the Circle.*

Henceforth we are concerned with the geometry in respect of a fundamental circle, of centre S and radius ρ. The points on the circle will be represented by binary symbols, the circular points being invariably represented by i, j. The types of (2, 1) correspondence on the circle which have a metrical significance are those which are specially related to ij. We consider two such types, namely:

(i) Auto-correspondences, or the (2, 1) correspondences which have the auto-pair ij.

(ii) Diametral correspondences, or the (2, 1) correspondences which have the apolar pair ij.

An auto-correspondence (or diametral correspondence) is completely determined by its triangle of fixed points $\alpha\beta\gamma$, and may be called "the auto-correspondence (or diametral correspondence) of $\alpha\beta\gamma$."

From the definition of correlate (16), it follows that the diametral correspondence of $\alpha\beta\gamma$ is the correlate of its auto-correspondence. The theorem 3·31 might therefore be utilized to derive the former from the latter.

* 4·10 shows that the *third* absolute parameters of the singular pencil $[P_1, P_1']$, and the pencil containing a perfect cube, are 0 and 8 respectively. It will be an excellent verification of 4·10 to obtain these values directly. The former value can be obtained by the method of limits, the latter by considering the pencil

$$\lambda\{(\beta x)^3 + 3(\beta x)^2(\gamma x)\} + \mu\{3(\beta x)(\gamma x)^2 + (\gamma x)^3\}$$

in respect of the representation $[P_1, P_2]$, where

$$P_1 = (\beta x)(\gamma x), \qquad P_2 = 3(\beta x)^2 + 2(\beta x)(\gamma x) + 3(\gamma x)^2.$$

Digitized by Google

A pair of points on the circle may be named indifferently by the chord joining them, or by the pole of this chord in respect of the circle. If two pairs on the circle be harmonic, the line which names either is incident with the point which names the other. Thus if the apolar pair of a (2, 1) correspondence F on the circle be named by the point O, then the pairs of the basic involution of F are the extremities of chords passing through O. F itself may accordingly be considered as a (1, 1) correspondence $\{F\}$ between points of the circle and lines through O. In particular, if F is a diametral correspondence, then O coincides with the centre S of the circle; thus the general diametral correspondence is a (1, 1) correspondence between diameters and points on the circle.

If F is the auto-correspondence of $\alpha\beta\gamma$, and if $y = p$ corresponds in F to $x = q_1, q_2$, the chord Oq_1q_2 may be called for brevity *the auto-chord* of p in respect of $\alpha\beta\gamma$. The (1, 1) correspondence $\{F\}$ is then the correspondence between points on the circle and their auto-chords. Since i, j correspond to each other in the auto-correspondence F, the auto-chords of each of i, j passes through the other. If now Oq, Oq' be the respective auto-chords of p, p', it must therefore follow from the property of a (1, 1) correspondence, that the cross ratios $O\{qq'ji\}$ and $\{pp'ij\}$ are equal. Hence follows the fundamental angle-property of the auto-correspondence:

5·1 *If the point p moves on the circle through an angle ω about the centre, its auto-chord turns through an angle* $-\frac{\omega}{2}$.

If the triangle $\alpha\beta\gamma$ is equilateral, then ij would be the Hessian pair of $\alpha\beta\gamma$, and the polar correspondence which has the fixed points $\alpha\beta\gamma$ would have ij for its apolar as well as its auto-pair (cf. 3·15). Thus:

5·2 *The diametral correspondence and the auto-correspondence of an equilateral triangle are both identical with its polar correspondence.*

We shall hereafter concern ourselves only with the auto-correspondence, the diametral correspondence being of little importance for our purpose.

The apolar cubic $f''(x)$ of the auto-correspondence F of $\alpha\beta\gamma$ has ij for its Hessian pair (3·20), and therefore represents the vertices $\alpha''\beta''\gamma''$ of a certain equilateral triangle. If $\alpha'\beta'\gamma'$ is the unique equilateral triangle apolar to $\alpha\beta\gamma$, then by 3·23, $\alpha''\beta''\gamma''$, $\alpha'\beta'\gamma'$, are cubicovariant equilateral triangles, so that $\alpha'\alpha''$, $\beta'\beta''$, $\gamma'\gamma''$ are diameters. Hence if $\alpha''\beta''\gamma''$ is given, the possible triangles $\alpha\beta\gamma$ are apolar to a fixed equilateral triangle $\alpha'\beta'\gamma'$, which is the cubicovariant triangle of $\alpha''\beta''\gamma''$. Also since similar correspondences have the same auto-pair (14), the correspondences similar to an auto-correspondence must also be auto-correspondences. It is therefore clear that:

Digitized by Google

5·3 *The auto-correspondences of the* ∞^2 *triangles apolar to an equilateral triangle constitute a family of similar correspondences.*

But if $F(x, y)$ is an auto-form (that is, a form which determines an auto-correspondence), the family of forms similar to F is given by $\mu F(x, y) + (ix)(jx)(\lambda y)$ (3·22). Thus the auto-chords of a point y, in regard to two similar auto-correspondences, are represented by two quadratics which differ by a multiple of $(ix)(jx)$, and are therefore parallel. Hence, using 5·3, we have the result:

5·4 *The auto-chords of a fixed point* y *in respect of all triangles apolar to an equilateral triangle, are parallel straight lines.*

(22) *The Pedal Angle.*

The fundamental angle-property 5·1 may be stated in the following symmetrical form:

5·5 *If the auto-chord in respect of* $\alpha\beta\gamma$, *of* one *vertex of any inscribed triangle* abc *is parallel to its opposite side, the same is true of every vertex; and the triangles* abc *with this property are all the triangles apolar to the equilateral triangle* $\alpha''\beta''\gamma''$ *of* (21).

The first part follows from 5·1; to prove the latter part of the theorem, we notice that the triangles abc are determined by any two of their vertices, and constitute therefore a linear ∞^2-system, apolar to a certain triangle. Now, if p_1p_2 is the auto-chord of a point p in respect of $\alpha\beta\gamma$, the triangles pp_1p_2 are of the type abc, and are apolar to $\alpha''\beta''\gamma''$ from the definition of the apolar triad (14). Hence the system of triangles abc is the system apolar to $\alpha''\beta''\gamma''$.

Let $\alpha\beta\gamma$, abc be two inscribed triangles, and let the auto-chords of a point p in respect of them be inclined at an angle θ with one another. Then by 5·1, the auto-chords of any other point p' in respect of the same triangles will also make the same angle θ with one another. Thus the angle θ depends only on the two triangles, and may be called their *pedal angle*. The triangles $\alpha\beta\gamma$, abc will also be said to be *pedo-parallel* when their pedal angle $\theta = 0 \pmod{\pi}$, and *pedo-perpendicular* when the pedal angle $= \frac{\pi}{2} \pmod{\pi}$. The ∞^2-system of all triangles pedo-parallel to $\alpha\beta\gamma$ may be called *the pedo-parallel system* of $\alpha\beta\gamma$. A pedo-parallel system is clearly a linear system, and it follows from 5·4 that:

5·6 *A pedo-parallel system may also be defined as the system of triangles apolar to an equilateral triangle.*

Thus the ∞^3 inscribed triangles of the circle fall into ∞^1 pedo-parallel systems, each of which is specified by its equilateral member, or by any one of its members. The triangles of the type

Digitized by Google

$ij\lambda$, being apolar to every equilateral triangle, belong to every pedo-parallel system.

5·7 *Cubicovariant equilateral triangles are pedo-perpendicular.*

For, let abc, $a'b'c'$ be cubicovariant equilateral triangles, so that aa', bb', cc' are diameters. By 5·2, the auto-chord of a in respect of abc and of a' in respect of $a'b'c'$, is in each case the diameter aa'. Hence from the fundamental angle-property 5·1, it follows that abc, $a'b'c'$ are pedo-perpendicular.

We can conclude from this theorem that the triangle $\alpha''\beta''\gamma''$, which represents the apolar triad of the auto-correspondence of $\alpha\beta\gamma$, is pedo-perpendicular to $\alpha\beta\gamma$. For the equilateral triangle $\alpha'\beta'\gamma'$, apolar to $\alpha\beta\gamma$, is also the pedo-parallel equilateral triangle of $\alpha\beta\gamma$ (5·6); and, from the last subsection (21), $\alpha''\beta''\gamma''$ is the cubicovariant equilateral triangle of $\alpha'\beta'\gamma'$. Hence $\alpha''\beta''\gamma''$ is the equilateral triangle pedo-perpendicular to $\alpha'\beta'\gamma'$, and therefore to $\alpha\beta\gamma$. It also follows that the triangles abc in theorem 5·5 are all the triangles pedo-perpendicular to $\alpha\beta\gamma$. This theorem might also be extended in the following manner:

5·8 *The auto-chords, in respect of $\alpha\beta\gamma$, of the vertices of any other inscribed triangle abc, make the same angle θ with their opposite sides; this θ is the complement of the pedal angle between $\alpha\beta\gamma$ and abc.*

The first part follows from 5·1. To prove the second part, we have only to take a triangle $\alpha_1\beta_1\gamma_1$, which makes the pedal angle $-\theta$ with $\alpha\beta\gamma$. Then abc is such that the auto-chords of its vertices, in respect of $\alpha_1\beta_1\gamma_1$, are parallel to their opposite sides; hence in virtue of our previous remark, abc is pedo-perpendicular to $\alpha_1\beta_1\gamma_1$, and therefore makes the pedal angle $\frac{\pi}{2} - \theta$ with $\alpha\beta\gamma$.

A special case of 5·8 arises if abc is identical with $\alpha\beta\gamma$. Since $\alpha\beta\gamma$ is pedo-parallel to itself, it follows that the auto-chords of α, β, γ, in respect of $\alpha\beta\gamma$, are perpendicular to their opposite sides. But these auto-chords should pass through α, β, γ respectively, and also through the point O which represents the apolar pair of the auto-correspondence of $\alpha\beta\gamma$. Hence we have the fundamental result:

5·9 *The perpendiculars of the triangle $\alpha\beta\gamma$ are concurrent; and the point of concurrence, namely the orthocentre, is none other than the point O which names the apolar pair of the auto-correspondence of $\alpha\beta\gamma$**.

* This might be verified algebraically. The canonical form 1·2 of the auto-form of $\alpha\beta\gamma$ may be shown to be

$$\Sigma\,(\alpha i)\,(\alpha j)\,(\beta\gamma)\,(\beta\gamma, ij)\,(\beta x)\,(\gamma x)\,(\alpha y),$$

where $(\beta\gamma, ij)$ is short for $(\beta i)\,(\gamma j) + (\beta j)\,(\gamma i)$.

By 3·6 the apolar quadratic of this is

$$\Sigma\,\frac{(\beta\gamma)\,(\alpha x)^2}{(\beta\gamma, ij)\,(\alpha i)\,(\alpha j)}.$$

Digitized by Google

If now p be any point of the circle, and pq the chord perpendicular to $\beta\gamma$, then in the triangle αpq the auto-chord of α being perpendicular to $\beta\gamma$, is parallel to pq. Therefore by 5·5 the auto-chord of p is parallel to αq. But this is a well-known property of the pedal line p. Hence:

5·10 *The auto-chord of p and the pedal line of p, in respect of the same triangle, are parallel to one another.*

This parallelism justifies the name "pedal angle." The following direct expression, which is easily obtained for the pedal angle between two triangles, might also have been made the starting point of a treatment of pedo-parallelism:

5·11 *The pedal angle between two inscribed triangles $\alpha\beta\gamma$, abc is equal to $\frac{1}{2}(a+b+c-\alpha-\beta-\gamma)$, where $\alpha, \beta, \gamma, a, b, c$ also represent he angles which the central radii vectores, to the corresponding points, make with a fixed direction.*

We may now set down the geometrical implications which certain of our previous results possess in virtue of 5·9. Thus the theorems 3·17, 3·18 of (13) give the following:

5.12 *The orthocentres of all inscribed triangles are definite with the exception of triangles of the types iiA, jjB, $ij\lambda$; the orthocentre of iiA is an indefinite point on iA, but assumes a definite limiting position if iiA is approached along any pencil Γ, this limiting position depending only on the linear system containing Γ and the singular pencil (iiA); the orthocentre of $ij\lambda$ is an indefinite point at infinity, but has a definite limiting position for any linear approach to $ij\lambda$, and the limiting position is the same for all approaches in the same pedo-parallel system.*

It is easy to show that 3·19 reduces to the following:

5·13 *If $\alpha\alpha'$ be a diameter of the circle, the orthocentre of the inscribed point-triangle $\alpha\alpha\alpha$ is the reflection of α' in α.*

We shall close this subsection with some invariant expressions relating to the triangle $\alpha\beta\gamma$ and the circular points ij. Writing

$$f(x) = (\alpha x)(\beta x)(\gamma x); \quad \phi(x) = (ix)(jx),$$

it is clear that the cubic $f'(x)$ which gives the pedo-parallel equi-

Now, the condition of perpendicularity of two chords $\beta\gamma$, xy is

$$(i\beta)(i\gamma)(jx)(jy)+(j\beta)(j\gamma)(ix)(iy)=0.$$

Hence the extremity of the chord through α, perpendicular to $\beta\gamma$, is

$$(i\beta)(i\gamma)(j\alpha)\,j+(j\beta)(j\gamma)(i\alpha)\,i,$$

so that we have only to show that the above apolar quadratic of the auto-form is apolar to

$$(\alpha x)\{(i\beta)(i\gamma)(j\alpha)(jx)+(j\beta)(j\gamma)(i\alpha)(ix)\}.$$

This is easily seen to be the case.

Digitized by Google

lateral triangle of $\alpha\beta\gamma$ is the cubic apolar to f and ϕ. It is easy to show that:

5·14 $$f'(x) = (ij)^2 f(x) + 3\phi(x)\{\phi(x), f(x)\}^2.$$

The pedo-perpendicular equilateral triangle of $\alpha\beta\gamma$ being the cubicovariant of $f'(x)$, is given by $f''(x)$, the Jacobian of f' and ϕ. Since $\{\phi(\phi,f)^2, \phi\}^1 = \frac{1}{3}\phi\{(\phi,f)^2, \phi\}^1 = -\frac{1}{3}\phi(\phi^2, f)^3$, by Gordan's series, we have the following expression for $f''(x)$:

5·15 $$f''(x) = (f', \phi)^1 = (ij)^2 (f, \phi)^1 - \phi(\phi^2, f)^3.$$

The linear covariant $(\phi, f)^2$ represents a point p (which we shall call *the Euler point* of $\alpha\beta\gamma$), such that ijp is apolar to $\alpha\beta\gamma$. Now the condition of apolarity of two triads $a_1b_1c_1$, $a_2b_2c_2$ of a conic, can be shown to be that the conic inscribed to $a_1b_1c_1$, $a_2b_2c_2$ should touch the polar lines of each of these triangles in regard to the conic. Hence the apolarity of $\alpha\beta\gamma$ and ijp implies that the inscribed parabola of $\alpha\beta\gamma$, with focus at p, touches the polar line of ijp in regard to the circle. But this polar line is seen to be the perpendicular-bisector of Sp. Thus the mid-point of Sp lies on the tangent at the vertex of the inscribed parabola, which is also the pedal line of p. Since this pedal line also bisects the join of p to the orthocentre O, it must be parallel to SO. Thus:

5·16 *The Euler point* $(\phi, f)^2$ *is the point whose pedal line in respect of* $\alpha\beta\gamma$ *is parallel to the Euler line of* $\alpha\beta\gamma$ *(namely, the line* SO*); the point* $(\phi^2, f)^3$ *is the diametrically opposite point of* $(\phi, f)^2$*.**

(23) *A Polar Correspondence Related to the Auto-correspondence.*

Let p_1p_2 be the diameter perpendicular to the auto-chord of p in respect of $\alpha\beta\gamma$. Then by 5·2, p_1p_2 is the auto-chord of p in respect of the pedo-perpendicular equilateral triangle of $\alpha\beta\gamma$. On applying 5·8 to the triangle pp_1p_2, it appears that the pedal lines of p_1, p_2 are parallel to pp_1, pp_2 respectively. Hence:

5·17 *If through each point* p_1 *we draw the chord* p_1p *parallel to the pedal line of* p_1 *in respect of* $\alpha\beta\gamma$*, the correspondence* (p_1, p) *is the* (2, 1) *polar correspondence of the pedo-perpendicular equilateral triangle of* $\alpha\beta\gamma$*.*

Let O' be an arbitrary point in the plane, r_1r_2 the auto-chord of p in respect of $\alpha\beta\gamma$, and q_1q_2 the chord through O' parallel to r_1r_2. We see immediately that the correspondence (p, q_1q_2) is an auto-correspondence. If $\alpha'\beta'\gamma'$ be its fixed points (so that O' is the orthocentre of $\alpha'\beta'\gamma'$), then the auto-chords of p in respect of $\alpha\beta\gamma$, $\alpha'\beta'\gamma'$ are by construction parallel, or $\alpha'\beta'\gamma'$ is pedo-parallel to $\alpha\beta\gamma$. Thus:

5·18 *The lines through* α', β', γ' *parallel to their respective auto-*

* For these two linear covariants and a related property of the Euler line, see F. Morley, "Some Polar Constructions," *Math. Ann.* Bd. 51.

Digitized by Google

chords in respect of $\alpha\beta\gamma$, are concurrent only if $\alpha'\beta'\gamma'$ is pedo-parallel to $\alpha\beta\gamma$; the point of concurrence is therefore the orthocentre of $\alpha'\beta'\gamma'$.

The triangles f', f'' of 5·14, 5·15, occur in connection with the study of the inscribed parabolas of $\alpha\beta\gamma$. It is easy to see that:

5·19 *If p_1p_2, q_1q_2 be the diameters parallel and perpendicular to the axis of the inscribed parabola whose focus is p, the correspondences (p, p_1p_2), (p, q_1q_2) are the respective polar correspondences of f'' and f'. The triangles f', f'' therefore give the foci of inscribed parabolas, whose axes touch the circle, and pass through its centre respectively.*

We know from 5·12 that the orthocentre of the triangle ijp has a definite limiting position P for all approaches in the pedo-parallel system of $\alpha\beta\gamma$; it is clear from the above, that P is the point at infinity on the directrix of the inscribed parabola whose focus is p.

(24) *Pedo-Parallel and Orthocentric Pencils.*

A *pedo-parallel pencil* of inscribed triangles is a pencil, any two members of which are pedo-parallel, that is, a pencil wholly contained in a pedo-parallel system.

Let O be any point of the plane, and let S, O stand also for the quadratics which represent parametrically the point-pairs on the circle which are named by S, O respectively (so that $S = (ix)(jx)$). Then from the definition of the two-quadratic representation, and from 5·9, it follows that the pencil $[S, O]$ consists of triangles having their orthocentre at O. Thus the inscribed triangles with a given orthocentre O form a pencil, which we may call *the orthocentric pencil* of O.

It is clear from 4·2 that any pencil, among the four Jacobian points of which i and j are included, is an orthocentric pencil. Since the singular pencil $ij\lambda$ belongs to every pedo-parallel system, a pedo-parallel pencil must contain a member $ij\lambda$. Hence:

5·20 *An orthocentric pencil may also be defined as a pencil which contains two members of the form iiA, jjB, the common orthocentre being the intersection of iA, jB (4·3); a pedo-parallel pencil may be defined as one which has a member of the form $ij\lambda$.*

Hence:

5·21 *The necessary and sufficient condition for pedo-parallelism of two inscribed triangles, is that the conic inscribed to them be a parabola; the locus of orthocentres of the pedo-parallel pencil determined by them is then the directrix of the parabola.*

By supposing that the quadratic P_1 represents the circular points on the circle, the theorem 4·7 reduces to:

5·22 *The apolar pencil of the orthocentric pencil of O is the orthocentric pencil of another point in the diameter through O.*

Digitized by Google

By the same supposition the family of pencils studied in (19) becomes the family of orthocentric pencils of points on a diameter A_1A_2 (say). The involution I_1 then corresponds to the involution of pairs of points in A_1A_2 which are inverse in regard to the fundamental circle; by footnote, p. 248, § 4, the involution I_3 corresponds to the involution of points on A_1A_2 equidistant from the centre. Further, if the two points determining the null orthocentric pencils are at a distance $\mu\rho$ from S, then, by 4·9 and the same footnote, $\mu^2 = 3$. Thus the fixed-points of I_2 are at a distance $\sqrt{3}\rho$ from S, a result which is also directly evident from 4·8. We have thus:

5·23 *The locus of points determining the null orthocentric pencils is a concentric circle of radius* $\sqrt{3}\rho$*; inverse points in respect of this circle determine apolar orthocentric pencils.*

The result may be verified in a particular case. If α is a point on the circle, the orthocentric pencil of α is the pencil of right-angled triangles which have their right angle at α. The apolar pencil therefore contains the member $(\alpha x)^3$, and is further an orthocentric pencil (5·22). Now by 5·13 the orthocentre of $\alpha\alpha\alpha$ is the point β on $S\alpha$ produced, which is distant $3S\alpha$ from S. So the orthocentric pencils of α, β are apolar orthocentric pencils. The points α, β being evidently inverse in the concentric circle of radius $\sqrt{3}\rho$, the theorem 5·23 is verified.

(25) *Orthocentric Loci.*

A general linear ∞^2-system $[M]$ of inscribed triangles consists of triangles apolar to a fixed triangle $\alpha\beta\gamma$. If O be the orthocentre of $\alpha\beta\gamma$, and O' the inverse of O in respect of the concentric circle of radius $\sqrt{3}\rho$, then the orthocentric pencil of O' being apolar to $\alpha\beta\gamma$ (5·23), must be completely contained in $[M]$. It is clear that this is the only orthocentric pencil contained in $[M]$. Also, the system $[M]$ has, in general, only one member in common with the pencil of equilateral triangles. A very convenient way of specifying $[M]$ is by means of its equilateral member and its unique orthocentric pencil.

Let M_1 be the member of $[M]$, of the form $ij\lambda$, and M_2, M_3 the members of the form iiA, jjB. If iA, jB intersect in O, then the pencil $[M_2, M_3]$ is the orthocentric pencil of O (5·20), and is the unique orthocentric pencil contained in $[M]$. Let M be the general member of $[M]$, and m its orthocentre. The correspondence between M, m is in general (1, 1), the exceptions being covered by 5·12. Since $[M]$ is a general linear system, the approaches to M_1 contained in $[M]$ are not pedo-parallel, so that the orthocentre of M_1 is indefinite, and may be anywhere at infinity (5·12). Similarly, the orthocentres of M_2, M_3 are indefinite, and may be anywhere on

Digitized by Google

Oi, Oj respectively. Hence the relation between M and m is a Cremona quadratic transformation with the fundamental elements $M = M_1, M_2, M_3$; $m = O, ji$. Now in a Cremona quadratic transformation between two planes, a linear locus in either is transformed into a quadratic locus passing through the three fundamental points in the other. Hence:

5·24 *The locus of orthocentres of any pencil of* $[M]$ *is a conic through* Oij, *that is, a circle* (*diametrically opposite points of this circle may be verified to be the orthocentres of pedo-perpendicular members of the pencil*); *the circle degenerates into a point-circle, when the pencil is orthocentric, and into a line when it is a pedo-parallel pencil* (5·21). *The locus of the orthocentres of a conic-system of triangles containing three members of the type* M_1, M_2, M_3 *is a straight line.*

It should be observed, however, that if $[M]$ is a pedo-parallel system (or a system containing either of the singular pencils iiA, jjB), the correspondence between M, m is no longer a Cremona, but a linear correspondence.

Suppose now that the circles C, C' are the orthocentric loci of two apolar pencils Γ, Γ'. Let $\alpha\beta\gamma$ be any member of Γ, and O its orthocentre, so that O is a point on C. Let O' be the point inverse to O in regard to the concentric circle of radius $\sqrt{3}\rho$. Then the apolar system of $\alpha\beta\gamma$ contains Γ' and also the orthocentric pencil of O'. Hence Γ' contains a member whose orthocentre is O', so that O' is a point on C'. It follows that C and C' are inverse circles in regard to the concentric circle of radius $\sqrt{3}\rho$. Thus:

5·25 *The orthocentric loci of two apolar pencils are two circles mutually inverse in the second fundamental circle* (*by which we mean the concentric circle of radius* $\sqrt{3}\rho$); *therefore the orthocentric locus of a null pencil is a circle orthogonal to the second fundamental circle.*

The latter part of this theorem may be verified for a particular case. The simplest null pencils are those of the type $pp\lambda$, where p is a fixed point. Now it is easy to see that the orthocentre of the inscribed line-triangle $pp\lambda$ is the reflection in p of the extremity of the chord through p perpendicular to $p\lambda$. Thus the orthocentric locus of the null pencil $(pp\lambda)$ is the reflection of the fundamental circle in p. It is an easy verification to show that the second fundamental circle of radius $\sqrt{3}\rho$ is orthogonal to the reflection of the fundamental circle in any of its points.

As a particular case of 5·25, we see that the orthocentric locus of a null pedo-parallel pencil must be a diameter A_1A_2. Now there are two remarkable harmonic point-pairs on a diameter A_1A_2, namely, the pair of extremities A_1A_2, and the pair $S\infty$. Hence there are four remarkable members of the pedo-parallel null pencil, namely, those whose orthocentres are these four points. The

Digitized by Google

members with their orthocentres at S, ∞ are the equilateral member and the member of the form $ij\lambda$ respectively; those which have their orthocentres at A_1, A_2 are the two right-angled members which have their right angles at A_1, A_2, and which (since they are pedo-parallel) have, by 5·11, perpendicular diameters for their hypotenuses. Hence:

5·26 *A null pedo-parallel pencil contains two harmonic pairs of remarkable members; the first pair consists of the two right-angled triangles of the pencil, the second of the equilateral triangle, and the triangle of the form* $ij\lambda$. *The former have their right angles at diametrically opposite points, and have perpendicular diameters for their hypoteneuses.*

(26) *Other Systems of Inscribed Triangles. The Antipedal Angle.*

Let every inscribed triangle $\alpha\beta\gamma$ be represented by the point in three-dimensional space whose homogeneous coordinates are proportional to the coefficients of the binary cubic $(\alpha x)(\beta x)(\gamma x)$. The coalesced triangles of the type *ttt* will then correspond to the point T of a certain fundamental twisted cubic C; in particular the two triads *iii*, *jjj* will correspond to two points, I, J on C. The equilateral triangles will correspond to points on the chord IJ of C, and the singular pencil $ij\lambda$ to the axis IJ (that is, the intersection of the osculating planes at I, J to C). Pedo-parallel systems are represented by planes through the axis IJ, pedo-parallel pencils by lines meeting the axis IJ, and orthocentric pencils by lines meeting the tangents at I, J.

The system of inscribed triangles congruent to a given inscribed triangle $\alpha\beta\gamma$ may be called *the congruent system of* $\alpha\beta\gamma$; it is clear that it is represented by a twisted cubic C' passing through the point P which corresponds to $\alpha\beta\gamma$. It is easy to show that C' passes through the points I, J of C, and has also the same osculating tetrahedron at IJ as C (so that C, C' have the same tangents and the same osculating planes at I, J). One and only one twisted cubic of the type C' passes through each point of space. It will be noticed also that C itself is a particular case of C'.

We define *equilateral systems* to be the linear ∞^2-systems which contain all equilateral triangles. Thus the equilateral systems correspond to the planes through the chord IJ of C, and are therefore, in a sense, correlative to the pedo-parallel systems. The triangle apolar to all members of an equilateral system is clearly of the form $ij\lambda$. Hence, from the definition of the Euler point (22), it follows that:

5·27 *An equilateral system is the totality of triangles which have a given Euler point.*

Digitized by Google

Just as the *pedal angle* between two triangles is the measure of separation between the *pedo-parallel systems* containing them, so there is a second angle, namely, the *antipedal angle*, which is the measure of separation between the *equilateral systems* containing them. We define the antipedal angle between two triangles to be half the angle subtended at the centre by their Euler points.

It was shown in 5·16 that the pedal line of the Euler point is parallel to the Euler line of the triangle. If we call the angle between the Euler lines of two inscribed triangles their *Eulerian angle*, then from 5·1 it easily follows that the pedal, antipedal, and Eulerian angles are connected by the following relation:

5·28 (*Pedal angle*)—(*antipedal angle*) = *Eulerian angle.*

Since the Eulerian angle between two co-orthocentric triangles is zero, we see, as a particular case, that:

5·29 *The pedal angle between two co-orthocentric triangles is also their antipedal angle.*

The condition that the pencil determined by two triangles may contain an equilateral member, is that their antipedal angle be equal to zero (mod π).

Digitized by Google

Integral Invariants of the Affine Field. By Mr M. H. A. NEWMAN, St John's College.

[*Received* 5 May, *read* 26 July, 1926.]

In this paper the method of infinitesimal transformations of coordinates, used by WEYL to determine conditions that a function of the tensors g_{ik} and ϕ_i, and certain of their derivatives, should be a scalar density, is applied (with certain modifications so as to give tensor relations) to functions of Γ^i_{rs} and $\Gamma^i_{rs,a}$. It is known* that in order that such a function should be a scalar density it must be a homogeneous function, of degree $\frac{1}{2}n$, of R^i_{jrs}, and this must of course be deducible from the equations found by the infinitesimal transformations. In view of the part which these equations may play, as "equations of energy," etc., in purely affine field theories, it seems desirable that the connection should be explicitly shown, and this is done in § 3.

§ 1. *A variation principle.*

Let M_n be a manifold having an assigned affine structure ($\Gamma^i_{rs} = \Gamma^i_{sr}$), D a domain in M_n, and $\mathfrak{P}_i$, $\mathfrak{Q}^r_i$, $\mathfrak{S}^{rs}_i$, $\mathfrak{T}^{rst}_i$, tensor-density fields, of the rank shown by their suffixes, having continuous components in D; and suppose that whatever the domain Ω in D and the vector field ξ^i,

$$\int_\Omega [\xi^i \mathfrak{P}_i + (\xi^i \mathfrak{Q}^r_i)_r + (\xi^i \mathfrak{S}^{rs}_i)_{rs} + (\xi^i \mathfrak{T}^{rst}_i)_{rst}]\, dx = 0 \quad \ldots\ldots(1)$$
$$(dx = dx_1 dx_2 \ldots dx_n).$$

(The existence of all relevant differential coefficients is assumed.)

If the values of the covariant derivatives are written out, the expression on the left takes, after suitable rearrangement, the form

$$\int_\Omega \left[\xi^i X_i + \frac{\partial}{\partial x_r}(\xi^i Y^r_i) + \frac{\partial^2}{\partial x_r \partial x_s}(\xi^i Z^{rs}_i) + \frac{\partial^3}{\partial x_r \partial x_s \partial x_t}(\xi^i W^{rst}_i)\right] dx = 0$$
$$\ldots\ldots(2),$$

and it may be inferred, since Ω as well as ξ^i is arbitrary, that

$$X_i = Y^r_i = Z^{(rs)}_i = W^{(rst)}_i = 0\dagger.$$

* Cf. Weitzenböck, *Invariantentheorie*, p. 358.

† $A^{(rs)} = \frac{1}{2}(A^{rs} + A^{sr})$, $A^{(rst)} = \frac{1}{6}(A^{rst} + A^{rts} + A^{str} + A^{srt} + A^{trs} + A^{tsr})$ (after SCHOUTEN).

Digitized by Google

But the quantities X, Y, Z, W are not all tensors. To get tensor equations it is preferable to proceed slightly differently.

In the first place it is plain that W_i^{rst} is $\mathfrak{T}_i^{rst}$, so that we may infer

$$\mathfrak{T}_i^{(rst)} = 0 \qquad \ldots\ldots(a)$$

from (1). Now if $\mathfrak{a}^{rs}$ is any tensor density of rank $(2+0)$ it is easily verified that $(\mathfrak{a}^{rs})_{rs} = (\mathfrak{a}^{rs})_{sr} = (\mathfrak{a}^{rs})_{(rs)}$. Hence (writing for a moment t^{rst} for $\mathfrak{T}_i^{rst}\xi^i$),

$$\begin{aligned}(t^{rst})_{rst} &= (t^{rst})_{r\,(st)} \\ &= (t^{rst})_{s\,(rt)} - (R^t_{ars}t^{rsa})_t \ * \\ &= (t^{rst})_{t\,(rs)} - (R^t_{ars}t^{rsa})_t - (R^r_{ast}t^{ast})_r,\end{aligned}$$

and, adding the three equations,

$$(t^{rst})_{rst} = (t^{(rst)})_{rst} - \tfrac{2}{3}(R^t_{ars}t^{rsa})_t - \tfrac{1}{3}(R^r_{ast}t^{ast})_r.$$

Since $t^{(rst)} = 0$ by (a), (1) reduces to

$$\int_\Omega [\xi^i\mathfrak{P}_i + \{\xi^i(\mathfrak{Q}_i^r - \tfrac{2}{3}R^r_{ast}\mathfrak{T}_i^{sta} - \tfrac{1}{3}R^r_{ast}\mathfrak{T}_i^{ast})\}_r + (\xi^i\mathfrak{S}_i^{rs})_{rs}]\,dx = 0 \quad \ldots\ldots(3).$$

The term corresponding in this expression to Z_i^{rs} of equation (2) is $\mathfrak{S}_i^{rs}$, and therefore

$$\mathfrak{S}_i^{(rs)} = 0 \qquad \ldots\ldots(b).$$

Since $$(\xi^i\mathfrak{S}_i^{rs})_{rs} = (\xi^i\mathfrak{S}_i^{(rs)})_{rs} = 0,$$

and since, whatever $\mathfrak{a}^i$, $(\mathfrak{a}^i)_i = \partial\mathfrak{a}^i/\partial x_i$, it follows that

$$\mathfrak{Q}_i^r = \tfrac{2}{3}R^r_{ast}\mathfrak{T}_i^{sta} + \tfrac{1}{3}R^r_{ast}\mathfrak{T}_i^{ast} \qquad \ldots\ldots(c),$$

$$\mathfrak{P}_i = 0 \qquad \ldots\ldots(d).$$

Equations a, b, c, and d are the required tensor relations.

§ 2.

Let $\mathfrak{A}\,(y^i_{pq},\ z^j_{rst})$ be a function of the $\frac{1}{2}n^2(n+1)^2$ variables y^i_{pq}, z^j_{rst}, where the indices take all integral values from 1 to n, save that $p \leqslant q$ and $r \leqslant s$. We proceed to investigate the conditions that $\mathfrak{A}\,(\Gamma^i_{pq},\ \Gamma^i_{rs,t})$† should be a scalar density.

* $R^a_{\beta ik} = +\dfrac{\partial\Gamma^a_{\beta k}}{\partial x_i} - \ldots.$

† $\Gamma^i_{rs,t} = \dfrac{\partial\Gamma^i_{rs}}{\partial x_t}.$

Digitized by Google

Since $$\int_{\Omega} \mathfrak{A}\, dx$$

is to be invariant for all transformations of coordinates, its value must be unaltered, in particular, by the infinitesimal transformation

$$\bar{x}_i = x_i + \xi^i,$$

ξ^i being a small vector field. Thus

$$0 = \int_{\bar{\Omega}} \bar{\mathfrak{A}}\,(\bar{x})\, d\bar{x} - \int_{\Omega} \mathfrak{A}\,(x)\, dx$$

$$= \int_{\Omega} \frac{\partial}{\partial x_i}(\mathfrak{A}\,\xi^i)\, dx + \int_{\Omega} \delta^* \mathfrak{A}\, dx \qquad \text{.....(4)†,}$$

where $$\delta^* \mathfrak{A} = \bar{\mathfrak{A}}\,(x) - \mathfrak{A}\,(x).$$

If θ is a quantity of any character and $\bar{\delta}\theta = \bar{\theta}\,(\bar{x}) - \theta\,(x)$, then

$$\delta^* \theta = \bar{\delta}\theta - \xi^i \frac{\partial \theta}{\partial x_i}.$$

Let the increment of $\mathfrak{A}$ when the Γ^i_{rs} receive a *general* small increment be

$$\delta\mathfrak{A} = \mathfrak{A}^{rs}_i\, \delta\Gamma^i_{rs} + \frac{\partial}{\partial x_t}(\mathfrak{B}^{rst}_i\, \delta\Gamma^i_{rs,t}) \qquad \text{.....(5),}$$

where $\mathfrak{A}^{rs}_i$ and $\mathfrak{B}^{rst}_i$ are symmetrical in r and s.

If $\mathfrak{A}$ is a scalar density, $\mathfrak{A}^{rs}_i$ and $\mathfrak{B}^{rst}_i$ are tensor densities. For, $\delta\Gamma^i_{rs}$ being a tensor, if its components vanish at a certain point P in one coordinate system they vanish in all. When this is so $\delta\Gamma^i_{rs,t}$ has tensor character at P, but save for the symmetry condition its value is unrestrained. Since $\delta\mathfrak{A} = \mathfrak{B}^{rst}_i\, \delta\Gamma^i_{rst}$ in this case, it follows that $\mathfrak{B}^{rst}_i$, itself symmetrical in r and s, is a tensor density; and so, therefore, is $\mathfrak{A}^{rs}_i$.

The quantities Γ^i_{rs} are transformed by the rule:

$$\bar{\Gamma}^i_{rs} = \Gamma^{\alpha}_{\beta\gamma} \frac{\partial x_\beta}{\partial \bar{x}_r} \cdot \frac{\partial x_\gamma}{\partial \bar{x}_s} \cdot \frac{\partial \bar{x}_i}{\partial x_\alpha} - \frac{\partial^2 \bar{x}_i}{\partial x_\beta \partial x_\gamma} \cdot \frac{\partial x_\beta}{\partial \bar{x}_r} \cdot \frac{\partial x_\gamma}{\partial \bar{x}_s}.$$

When $\bar{x}_i = x_i + \xi^i$,

$$\frac{\partial \bar{x}_i}{\partial x_j} = \delta^i_j + \frac{\partial \xi^i}{\partial x_j}, \quad \frac{\partial x_i}{\partial \bar{x}_j} = \delta^i_j - \frac{\partial \xi^i}{\partial x_j},$$

† Cf. Weyl, *Raum, Zeit, Materie*, 5te Aufl., pp. 111, 309; also Weitzenböck, *op. cit.*, p. 372; Pauli, *Relativitätstheorie*, § 23.

Digitized by Google

and therefore

$$\bar{\delta}\Gamma^i_{rs} = -\Gamma^i_{\beta s}\frac{\partial \xi^\beta}{\partial x_r} - \Gamma^i_{r\gamma}\frac{\partial \xi^\gamma}{\partial x_s} + \Gamma^a_{rs}\frac{\partial \xi^i}{\partial x_a} - \frac{\partial^2 \xi^i}{\partial x_r \partial x_s},$$

and

$$\begin{aligned}\delta^* \Gamma^i_{rs} &= \bar{\delta}\Gamma^i_{rs} - \xi^a \frac{\partial \Gamma^i_{rs}}{\partial x_a} \\ &= -(\xi^i)_{rs} + \xi^a R^i_{rsa}.\end{aligned}$$

Substituting this in

$$\delta^* \mathfrak{A} = \mathfrak{A}^{rs}_i \delta^* \Gamma^i_{rs} + (\mathfrak{B}^{rst}_i \delta^* \Gamma^i_{rs})_t$$

we find for the expression on the right of (4):

$$\begin{aligned}\int_\Omega [\mathfrak{A}^{rs}_i \{-(\xi^i)_{rs} + \xi^a R^i_{rsa}\} + \{\mathfrak{A}\xi^t - \mathfrak{B}^{rst}_i (\xi^i)_{rs} + \mathfrak{B}^{rst}_i \xi^a R^i_{rsa}\}_t] dx \\ = \int_\Omega [\xi^i (-(\mathfrak{A}^{rs}_i)_{rs} + R^a_{rsi}\mathfrak{A}^{rs}_a) \\ + \{\xi^i (\mathfrak{A}\delta^t_i + 2(\mathfrak{A}^{st}_i)_s - (\mathfrak{B}^{rst}_i)_{rs} + \mathfrak{B}^{rst}_a R^a_{rsi})\}_t \\ + \{\xi^i (-\mathfrak{A}^{rs}_i + 2(\mathfrak{B}^{trs}_i)_t)\}_{rs} \\ - (\xi^i \mathfrak{B}^{rst}_i)_{rst}] dx.\end{aligned}$$

Hence the principle of § 1 applied to equation 4 gives

$$\mathfrak{B}^{(rst)}_i = 0 \qquad \text{.....(A)},$$

$$\mathfrak{A}^{rs}_i - 2(\mathfrak{B}^{t(rs)}_i)_t = 0 \qquad \text{.....(B)},$$

$$\mathfrak{A}\delta^t_i + 2(\mathfrak{A}^{st}_i)_s - (\mathfrak{B}^{rst}_i)_{rs} + R^a_{rsi}\mathfrak{B}^{rst}_a + \tfrac{1}{3} R^t_{ars}\mathfrak{B}^{ars}_i = 0 \quad \text{...(C)},$$

$$(\mathfrak{A}^{rs}_i)_{rs} - R^a_{rsi}\mathfrak{A}^{rs}_a = 0 \qquad \text{.....(D)}.$$

(A), (B), (C), and (D) are the necessary and sufficient conditions that $\mathfrak{A}$ should have scalar density character under all "deformations"—transformations generated by infinitesimal displacements.

§ 3.

From (A) and the fact that $\mathfrak{B}^{rst}_i = \mathfrak{B}^{srt}_i$ it follows that

$$\mathfrak{B}^{rst}_i = -2\mathfrak{B}^{t(rs)}_i \qquad \text{.....(6)},$$

and therefore (B) may be replaced by

$$\mathfrak{A}^{rs}_i = -(\mathfrak{B}^{rst}_i)_t \qquad \text{.....(B')}.$$

Digitized by Google

Now the necessary and sufficient condition that $\mathfrak{A}(y^i_{pq}, z^j_{rst})$, a function of the $\frac{n^2(n+1)^2}{2}$ variables y^i_{pq}, z^j_{rst}, should in fact be a function of the $\frac{n^2(n-1)^2}{3}$ distinct and non-zero functions x^i_{jkl}, where

$$x^i_{jkl} = z^i_{jlk} - z^i_{jkl} - y^a_{jk} y^i_{al} + y^a_{jl} y^i_{ak} \text{ *},$$

is that functions $\mathfrak{f}^{\beta\gamma\delta}_a$, antisymmetrical in γ and δ, exist, such that

$$\frac{\partial \mathfrak{A}}{\partial y^i_{rs}} = \mathfrak{f}^{\beta\gamma\delta}_a \cdot \frac{\partial x^a_{\beta\gamma\delta}}{\partial y^i_{rs}}, \quad \frac{\partial \mathfrak{A}}{\partial z^i_{rst}} = \mathfrak{f}^{\beta\gamma\delta}_a \cdot \frac{\partial x^a_{\beta\gamma\delta}}{\partial z^i_{rst}};$$

that is, such that

$$\frac{\partial \mathfrak{A}}{\partial y^i_{rs}} = -2(\mathfrak{f}^{(rs)\beta}_a y^a_{i\beta} + \mathfrak{f}^{\beta\gamma(r}_i y^{s)}_{\beta\gamma}) \qquad \text{.....(7)},$$

$$\frac{\partial \mathfrak{A}}{\partial z^i_{rst}} = -2\mathfrak{f}^{(rs)t}_i \qquad \text{.....(8)}.$$

These equations involve only the symmetrical part, $\mathfrak{f}^{(rs)t}_i$, of $\mathfrak{f}^{rst}_i$; and (8) may be regarded as giving the definition of this part and (7) the condition imposed on $\mathfrak{A}$. When $y^i_{pq} = \Gamma^i_{pq}$ and $z^j_{rst} = \Gamma^j_{rs,t}$, $\frac{\partial \mathfrak{A}}{\partial z^j_{rst}}$ is identical with $\mathfrak{B}^{rst}_i$ and $\frac{\partial \mathfrak{A}}{\partial y^i_{pq}}$ with $\mathfrak{A}^{pq}_i + \frac{\partial}{\partial x_t}(\mathfrak{B}^{pqt}_i)$. Hence condition (7) is

$$\begin{aligned}
\mathfrak{A}^{rs}_i &= -\frac{\partial}{\partial x_t}(\mathfrak{B}^{rst}_i) + \mathfrak{B}^{rst}_a \Gamma^a_{it} + \mathfrak{B}^{\beta\gamma(r}_i \Gamma^{s)}_{\beta\gamma} \\
&= -\frac{\partial}{\partial x_t}(\mathfrak{B}^{rst}_i) + \Gamma^a_{it} \mathfrak{B}^{rst}_a - \Gamma^r_{\beta t} \mathfrak{B}^{\beta st}_i - \Gamma^s_{\beta t} \mathfrak{B}^{r\beta t}_i \text{ (by 6)} \\
&= -(\mathfrak{B}^{rst}_i)_t,
\end{aligned}$$

equations (B′). The import of the equations (B′) is, then, that *all scalar densities which are functions of* Γ^i_{rs} *and* $\Gamma^i_{rs,t}$ *alone are functions of* R^i_{rst}. The functions $\mathfrak{f}^{rst}_i$ are equal to $\partial \mathfrak{A}/\partial R^i_{rst}$.

Equations (C) may be simplified by noticing that

$$\begin{aligned}
(\mathfrak{B}^{rst}_i)_{rs} &= -(\mathfrak{B}^{str}_i)_{rs} - (\mathfrak{B}^{trs}_i)_{rs} \\
&= (\mathfrak{A}^{st}_i)_s - (\mathfrak{B}^{str}_i)_{rs} - 2(\mathfrak{B}^{trs}_i)_{[rs]} \text{ †} \\
&= 2(\mathfrak{A}^{st}_i)_s + R^t_{ars} \mathfrak{B}^{ars}_i - R^a_{irs} \mathfrak{B}^{trs}_a,
\end{aligned}$$

* When $r > s$, y^i_{rs} and z^i_{rst} stand for y^i_{sr} and y^i_{srt}.

† $A^{[rs]} = \frac{1}{2}(A^{rs} - A^{sr})$ (after SCHOUTEN).

Digitized by Google

which combined with (C) gives

$$\mathfrak{A}\delta_i^t + R^a_{rsi}\,\mathfrak{B}_a^{rst} - \tfrac{2}{3}R^t_{ars}\,\mathfrak{B}_i^{ars} + R^a_{irs}\,\mathfrak{B}_a^{trs} = 0 \quad \ldots\ldots(C').$$

Now
$$\tfrac{1}{2}\mathfrak{f}_i^{rst}R^j_{rst} = \tfrac{1}{2}\mathfrak{f}_i^{rst}(-R^j_{trs} + R^j_{srt})$$
$$= \mathfrak{f}_i^{rst}R^j_{srt}, \quad (\text{since } \mathfrak{f}_i^{rst} = -\mathfrak{f}_i^{rts}),$$

whence
$$\tfrac{3}{2}\mathfrak{f}_i^{rst}R^j_{rst} = 2\mathfrak{f}_i^{(rs)t}R^j_{rst} = -\mathfrak{B}_i^{rst}R^j_{rst}.$$

Hence
$$0 = \mathfrak{A}\delta_i^t - 2R^a_{rsi}\mathfrak{f}_a^{rst} - R^a_{irs}\mathfrak{f}_a^{trs} + R^t_{ars}\mathfrak{f}_i^{ars} \quad \ldots\ldots(C'');$$

and, contracting,

$$n\mathfrak{A} = 2R^a_{rst}\mathfrak{f}_a^{rst} = 2R^a_{rst}\frac{\partial\mathfrak{A}}{\partial R^a_{rst}}.$$

Thus the proposition that $\mathfrak{A}$ *is a homogeneous function of* R^i_{rst}, *of degree* $n/2$, is equivalent to the contracted form of (C).

Again, since $\mathfrak{A}$ is a function of R^i_{rst},

$$\frac{\partial\mathfrak{A}}{\partial x_i} = \frac{\partial\mathfrak{A}}{\partial R^a_{\beta rs}}\cdot\frac{\partial R^a_{\beta rs}}{\partial x_i} = \mathfrak{f}_a^{\beta rs}\frac{\partial R^a_{\beta rs}}{\partial x_i} \quad \ldots\ldots(9).$$

If then it is known that $\mathfrak{A}$ is a tensor, the tensor equation

$$(\mathfrak{A})_i = \mathfrak{f}_a^{\beta rs}(R^a_{\beta rs})_i \quad \ldots\ldots(10),$$

which is true in geodesic coordinates, may be inferred to be true in all. We may expect that this will also follow purely algebraically from (A)—(D), and in fact (C″) is really a statement that (9) and (10) are equivalent. For

$$(\mathfrak{A})_i - \mathfrak{f}^{rst}(R^a_{rst})_i = \frac{\partial\mathfrak{A}}{\partial x_i} - \mathfrak{f}^{rst}\frac{\partial R^a_{rst}}{\partial x_i}$$
$$- \Gamma^a_{ia}\,\mathfrak{A} - \mathfrak{f}_a^{rst}(\Gamma^a_{\beta i}R^\beta_{rst} - \Gamma^\beta_{ri}R^a_{\beta st} - \Gamma^\beta_{si}R^a_{r\beta t} - \Gamma^\beta_{ti}R^a_{rs\beta})$$
$$= \frac{\partial\mathfrak{A}}{\partial x_i} - \mathfrak{f}^{rst}\frac{\partial R^a_{rst}}{\partial x_i} - \Gamma^a_{\beta i}(\mathfrak{A}\delta^\beta_a + \mathfrak{f}_a^{rst}R^\beta_{rst} - \mathfrak{f}_r^{\beta st}R^r_{ast} - 2\mathfrak{f}_t^{rs\beta}R^t_{rsa}),$$

and the expression in the bracket is the right-hand side of (C″).

The four equations (D) are among those found by a general "minimum principle"

$$\delta\int\mathfrak{A}\,dx = 0,$$

and form the "equations of energy" for field theories derived from such a law. From the interpretation that has now been given it is plain that (A), (B), and (C) are mutually independent. Equations (D), on the other hand, can be derived from them, thus:

Digitized by Google

Taking the "divergence" of (C″) we have

$$2(R^a_{rsi}\mathfrak{f}^{rst}_a)_t + (R^a_{irs}\mathfrak{f}^{trs}_a)_t - (R^t_{ars}\mathfrak{f}^{ars}_i)_t = (\mathfrak{A})_i = \mathfrak{f}^{rst}_a (R^a_{rst})_i \quad \ldots(11).$$

Since $\mathfrak{f}^{rst}_a$ is antisymmetrical in s and t,

$$0 = \mathfrak{f}^{rst}_a [(R^a_{rsi})_t + (R^a_{rti})_s]$$

$$= \mathfrak{f}^{rst}_a [-(R^a_{rts})_i - (R^a_{rit})_s + (R^a_{rti})_s], \quad \text{(Bianchi's identity)},$$

whence $$\mathfrak{f}^{rst}_a (R^a_{rst})_i = 2\mathfrak{f}^{rst}_a (R^a_{rsi})_t.$$

Substituting in (11),

$$2(\mathfrak{f}^{rst}_a)_t R^a_{rsi} + (R^a_{irs}\mathfrak{f}^{trs}_a)_t - (R^t_{ars}\mathfrak{f}^{ars}_i)_t = 0 \quad \ldots\ldots(12).$$

From equation (B′),

$$(\mathfrak{A}^{rs}_i)_{rs} = 2(\mathfrak{f}^{rst}_i)_{t\,(rs)}$$

$$= 2(\mathfrak{f}^{rst}_i)_{tsr} + R^a_{irs}(\mathfrak{f}^{rst}_a)_t$$

$$= 2(\mathfrak{f}^{rst}_i)_{[ts]r} + R^a_{irs}(\mathfrak{f}^{rst}_a)_t$$

$$= (R^r_{ast}\mathfrak{f}^{ast}_i)_r - (R^a_{ist}\mathfrak{f}^{rst}_a)_r + R^a_{irs}(\mathfrak{f}^{rst}_a)_t$$

$$= (\mathfrak{f}^{rst}_a)_t (2R^a_{rsi} + R^a_{irs}), \text{ by (12)}$$

$$= 2(\mathfrak{f}^{rst}_a)_t R^a_{(rs)\,i} = \mathfrak{A}^{rs}_a R^a_{rsi}.$$

If $\mathfrak{A}$ is a function of R_{ik} only, and $\partial\mathfrak{A}/\partial R_{ik} = \mathfrak{h}^{ik}$, it is easily verified* that

$$\mathfrak{A}^{rs}_i = (\mathfrak{h}^{(rs)})_i - \delta^{(r}_i (\mathfrak{h}^{s)\,t})_t,$$

and that $$\mathfrak{B}^{rst}_i = \delta^t_i \mathfrak{h}^{(rs)} - \delta^{(r}_i \mathfrak{h}^{s)\,t}.$$

Equations (C) reduce in this case to

$$\mathfrak{A}\delta^t_i = R_{ir}\mathfrak{h}^{tr} + R_{ri}\mathfrak{h}^{rt}.$$

* Cf. Einstein, "Zur allgemeinen Relativitätstheorie," *Sitzungsber. der Preuss. Akad.*, 1923, p. 35.

Digitized by Google

An approximation to the motion of two rotating electrical doublets in a plane. By P. A. TAYLOR, B.A., Emmanuel College.

[*Received* 22 May, *read* 26 July, 1926.]

§ 1. *Introduction.* This problem was suggested by an article of Debye's* in which he considers the effect on a free electron of an electrical doublet constrained to rotate uniformly about a line through itself perpendicular to its own axis. Assuming that the free electron is initially at rest, he shows that to a first approximation it performs simple harmonic vibrations about its initial position. Proceeding to a second approximation, he shows that the free electron appears to be repelled by the rotating doublet. In this paper, instead of considering an electron and a rotating doublet, we consider two rotating doublets. In order to simplify the calculations we suppose that the two doublets are constrained to lie in a plane, but are otherwise free. The principal results of the calculations will be found collected in the summary at the end.

The problem is of interest as it might give a clue to the origin of forces between atoms. For Debye shows that the field due to a hydrogen atom at any distance greater than about ten times its radius is to a close approximation the same as the field of the equivalent rotating doublet†. We shall find that two rotating doublets behave as if they acted on each other with a force of the same general type as the interatomic forces found empirically by Lennard-Jones and the present author‡. Comparison can only be made as regards order of magnitude for three reasons. We have constrained our doublets to lie in a plane. Secondly, we shall consider the moment of inertia and the electrical moment of a doublet as constant, whereas in an encounter between two hydrogen atoms the radius vector from the proton to the electron of the same atom might suffer perturbations, so that the moment of inertia and the electrical moment may be variable. Thirdly, no gas so simple as a gas of hydrogen atoms exists.

§ 2. The plane of the paper is the plane in which the doublets are constrained to lie.

G is the centre of gravity of the two doublets O_1 and O_2.

* P. Debye, "Molekularkräfte und ihre elektrische Deutung, § 6," *Phys. Zeit.*, vol. XXII, p. 306 (1921).

† P. Debye, *loc. cit.*

‡ J. E. Jones, *Proc. Roy. Soc.* A, vol. CVI, pp. 441, 463, 709 (1924); vol. CVII, p. 157 (1925); J. E. Lennard-Jones and P. A. Taylor, *Proc. Roy. Soc.* A, vol. CIX, p. 476 (1925).

Digitized by Google

m_1 and m_2 are the masses of the two doublets O_1 and O_2 respectively.

μ_1 and μ_2 are the electrical moments of the two doublets O_1 and O_2 respectively.

I_1 and I_2 are the moments of inertia of the two doublets O_1 and O_2 respectively.

Gx and Gy are rectangular axes through G parallel to directions fixed in space.

The rest of the notation is as in the figure.

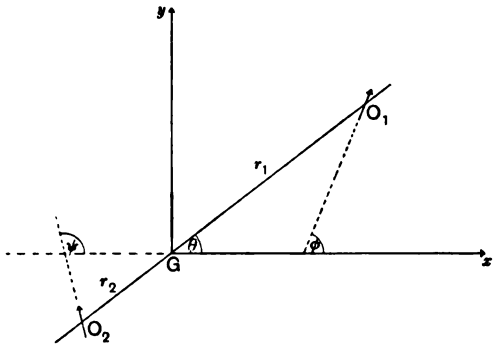


Fig. 1

We write also $\epsilon = \psi - \phi,$

$$r = r_1 + r_2.$$

Then, if T is the kinetic energy of the motion relative to G, and if V is the potential energy, we find

$$2T = M(\dot{r}^2 + r^2\dot{\theta}^2) + I_1\dot{\phi}^2 + I_2\dot{\psi}^2 \qquad \text{.....(2·0),}$$

$$2V = \frac{2\mu_1\mu_2}{r^3}\{\cos\epsilon - 3\cos(\phi - \theta)\cos(\psi - \theta)\}$$

$$= -\frac{\mu_1\mu_2}{r^3}\{3\cos(\phi + \psi - 2\theta) + \cos(\phi - \psi)\} \qquad \text{...(2·1),}$$

where $$M = \frac{m_1 m_2}{m_1 + m_2} \qquad \text{.....(2·2),}$$

and we have at once Lagrange's equations of motion.

Suppose now that the doublets are at such a great distance apart that the coupling (i.e. the influence of one doublet on the motion of the other) is small. If there were no coupling each doublet would move with uniform velocity in a straight line and would

Digitized by Google

rotate with constant angular velocity. Thus, neglecting the coupling, we have as an approximate solution

$$\left.\begin{aligned} r\cos\theta &= ut + \alpha \\ r\sin\theta &= vt + \beta \\ \phi &= \phi_1 t + \phi_0 \\ \psi &= \psi_1 t + \psi_0 \end{aligned}\right\} \qquad \text{.....(2·3).}$$

The arbitrary constants $\alpha, \beta, \phi_0, \psi_0, u, v, \phi_1, \psi_1$ may be so chosen that the equations (2·3) determine the exact values of $r, \theta, \phi, \psi, \dot{r}, \dot{\theta}, \dot{\phi}, \dot{\psi}$ at any given instant of time (say $t = 0$). Now let u, v, ϕ_1, ψ_1 be chosen in this manner, i.e. the vector $[u, v]$ is the relative linear velocity of the two doublets, and ϕ_1 and ψ_1 the angular velocities of the doublets at time $t = 0$. Then $\alpha, \beta, \phi_0, \psi_0$ may be taken as new variables, defined in terms of r, θ, ϕ, ψ, t by equations (2·3). We have seen that an approximate solution of the equations of motion is given by $\alpha = \text{const.}, \beta = \text{const.}, \phi_0 = \text{const.}, \psi_0 = \text{const.}$ Therefore in the exact solution $\dot{\alpha}, \dot{\beta}, \dot{\phi}_0, \dot{\psi}_0$ will be small, and so also will be $\ddot{\alpha}, \ddot{\beta}, \ddot{\phi}_0, \ddot{\psi}_0$. (Also from the way in which u, v, ϕ_0, ψ_0 were chosen, at time $t = 0$ we have $\dot{\alpha} = 0, \dot{\beta} = 0, \dot{\phi}_0 = 0, \dot{\psi}_0 = 0$ accurately.)

We may express the equations of motion after some algebraic reduction in the following form:

$$\left.\begin{aligned} \ddot{\alpha} &= A\cos(\phi_1+\psi_1)t + B\sin(\phi_1+\psi_1)t \\ &\quad + C\cos(\phi_1-\psi_1)t + D\sin(\phi_1-\psi_1)t \\ \ddot{\beta} &= A'\cos(\phi_1+\psi_1)t + B'\sin(\phi_1+\psi_1)t + \dots \\ \ddot{\phi} &= A''\cos(\phi_1+\psi_1)t + \dots \\ \ddot{\psi} &= A'''\cos(\phi_1+\psi_1)t + \dots \end{aligned}\right\} \qquad (2\cdot4).$$

The A's, B's, etc. are expressions of which the following are typical:

$$\left.\begin{aligned} A &= -\frac{15\mu_1\mu_2}{4Mr^4}\cos(\phi_0+\psi_0-3\theta) - \frac{3\mu_1\mu_2}{4Mr^4}\cos(\phi_0+\psi_0-\theta) \\ B &= \frac{15\mu_1\mu_2}{4Mr^4}\sin(\phi_0+\psi_0-3\theta) + \frac{3\mu_1\mu_2}{4Mr^4}\sin(\phi_0+\psi_0-\theta) \\ C'' &= -\frac{\mu_1\mu_2}{2I_1r^3}\sin(\phi_0-\psi_0) \\ D'' &= -\frac{\mu_1\mu_2}{2I_1r^3}\cos(\phi_0-\psi_0) \end{aligned}\right\} \qquad (2\cdot5).$$

The right-hand sides of equations (2·4) are functions of t, both directly and *via* the A's, B's, C's, D's, which are given as functions of $r, \theta, \phi_0, \psi_0$ by equations (2·5) and other equations of similar form.

Digitized by Google

The two variables r and θ are given by equations (2·3) as functions of t, both directly and *via* α and β. To a first approximation we may neglect the variation of the four quantities α, β, ϕ_0, ψ_0, as we have shown that they are approximately constant. If the angular velocities of the doublets are large compared with the ratio of their linear relative velocity to their distance apart, we can consider r and θ as approximately constant. Thus for a first approximation we can consider the A's, B's etc. as constant, i.e. their variations are slow compared with the variations of

$$\begin{matrix}\cos\\ \sin\end{matrix}(\phi_1 \pm \psi_1)\, t.$$

We can now integrate equations (2·4) and obtain equations like

$$\begin{aligned}\alpha - \alpha_0 = {} & \frac{A}{(\phi_1+\psi_1)^2}[1-\cos(\phi_1+\psi_1)t] \\ & + \frac{B}{(\phi_1+\psi_1)^2}[(\phi_1+\psi_1)t - \sin(\phi_1+\psi_1)t] \\ & + \frac{C}{(\phi_1-\psi_1)^2}[1-\cos(\phi_1-\psi_1)t] \\ & + \frac{D}{(\phi_1-\psi_1)^2}[(\phi_1-\psi_1)t - \sin(\phi_1-\psi_1)t] \quad \ldots(2\cdot6),\end{aligned}$$

where α_0, β_0, ϕ_{00}, ψ_{00} are the values of α, β, ϕ_0, ψ_0 respectively when $t = 0$.

For a second approximation we must no longer consider the A's, B's, etc. as constants. Let r_0, r_1, θ_0, θ_1 be the values at time $t = 0$ of r, $\dot{r}$, θ, $\dot{\theta}$ respectively. Then since we have already assumed that the linear relative velocity of the doublets is small, we may write, for a few periods of revolution of the doublets at least

$$\begin{aligned} r &= r_0 + r_1 t, \\ \theta &= \theta_0 + \theta_1 t. \end{aligned}$$

Expanding the A's etc. to the first order in r_1, θ_1, and also in $(\alpha - \alpha_0)$, $(\beta - \beta_0)$, etc., we have (letting the suffix 0 denote the values at time $t = 0$)

$$\begin{aligned} A = {} & A_0 + \left(\frac{\partial A}{\partial r}\right)_0 \dot{r}_1 t + \left(\frac{\partial A}{\partial \theta}\right)_0 \theta_1 t + \left(\frac{\partial A}{\partial \alpha}\right)_0 (\alpha - \alpha_0) + \left(\frac{\partial A}{\partial \beta}\right)_0 (\beta - \beta_0) \\ & + \left(\frac{\partial A}{\partial \phi_0}\right)_0 (\phi_0 - \phi_{00}) + \left(\frac{\partial A}{\partial \psi_0}\right)_0 (\psi_0 - \psi_{00}) \quad (2\cdot70) \\ = {} & E_a + F_a t + G_a \cos(\phi_1+\psi_1)t + H_a \sin(\phi_1+\psi_1)t \\ & + K_a \cos(\phi_1-\psi_1)t + L_a \sin(\phi_1-\psi_1)t \\ & \qquad \ldots\ldots(2\cdot71). \end{aligned}$$

Digitized by Google

The E_a's, F_a's, etc. are constants whose values may be found by substituting in equation (2·70) from (2·6) and similar equations and comparing the result with equation (2·71). Similarly

$$B' = E'_b + F'_b t + \ldots.$$

Substituting in equations (2·4) from equations (2·71) we have

$$\begin{aligned}\ddot{\alpha} = \{&E_a + F_a t + G_a \cos(\phi_1+\psi_1)t + H_a \sin(\phi_1+\psi_1)t + K_a \cos(\phi_1-\psi_1)t \\ &+ L_a \sin(\phi_1-\psi_1)t\} \cos(\phi_1+\psi_1)t \\ + \{&E_b + F_b t + \ldots \quad + L_b \sin(\phi_1-\psi_1)t\} \sin(\phi_1+\psi_1)t \\ + \{&E_c + F_c t + \ldots \quad + L_c \sin(\phi_1-\psi_1)t\} \cos(\phi_1-\psi_1)t \\ + \{&E_d + F_d t + \ldots \quad + L_d \sin(\phi_1-\psi_1)t\} \sin(\phi_1-\psi_1)t \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \ldots\ldots(2\cdot8).\end{aligned}$$

§ 3. We will first suppose that neither $(\phi_1 + \psi_1)$ nor $(\phi_1 - \psi_1)$ is small, i.e. $(\phi_1 + \psi_1)$ and $(\phi_1 - \psi_1)$ are to be considered as large and of the same order of magnitude as ϕ_1 and ψ_1. We will find the average relative acceleration of the doublets over a time $t = -\tau$ to $t = +\tau$, where the time 2τ is sufficiently large to contain a large number of periods of

$$\begin{matrix}\cos \\ \sin\end{matrix} (\phi_1 + \psi_1)\, t \text{ and } \begin{matrix}\cos \\ \sin\end{matrix} (\phi_1 - \psi_1)\, t,$$

and yet sufficiently small for r and θ to be effectively constant*. Letting a bar denote the average value with respect to t over the time $-\tau$ to $+\tau$ we have

$$\overline{t \cos \omega t} = 0,$$

where ω denotes either $(\phi_1 + \psi_1)$ or $(\phi_1 - \psi_1)$.

$$\overline{t \sin \omega t} = -\frac{1}{\omega} \cos \omega\tau + \frac{1}{\tau\omega^2} \sin \omega\tau.$$

The first term in this expression fluctuates, going through a cycle of values, equally positive and negative, in the time of one period of $\begin{matrix}\cos \\ \sin\end{matrix} \omega t$. We are trying to compare the actual motion with one in which the effects of individual rotations are smoothed out; and this term $(-\cos \omega\tau/\omega)$ contributes nothing permanent to the

* This is possible as regards $(\phi_1+\psi_1)\,t$, but not as regards $(\phi_1-\psi_1)\,t$, when the doublets represent normal Bohr hydrogen atoms, rotating in the same sense, and whose relative velocity is the kinetic theory velocity appropriate to a temperature of 0°C. For then the distance travelled in 30 periods of revolution of the doublets is about $\frac{1}{10}$ of the radius of a hydrogen atom, and hence a considerably smaller fraction of the distance apart of the two doublets. To make a suitable choice of τ possible as regards both $(\phi_1+\psi_1)\,t$ and $(\phi_1-\psi_1)\,t$ we should have to make our doublets correspond to hydrogen atoms with different quantum numbers.

Digitized by Google

averaged acceleration. Besides, owing to the factor $1/\omega$, the extreme values of $(-\cos\omega\tau/\omega)$ would be small compared with $\overline{\cos^2\omega t}$ (i.e. $\frac{1}{2}$). A fortiori, the term $(\sin\omega\tau/\tau\omega^2)$, which also fluctuates through positive and negative values, tends to zero as τ increases, and contains a factor $1/\omega^2$, contributes nothing permanent to the averaged acceleration. So we write

$$\overline{t\sin\omega t} = 0.$$

Similarly

$$\overline{\begin{matrix}\cos\\ \sin\end{matrix}\omega t} = 0, \qquad \overline{\cos\omega t\sin\omega t} = 0, \qquad \overline{\begin{matrix}\cos^2\\ \sin^2\end{matrix}\omega t} = \tfrac{1}{2},$$

$$\overline{\begin{matrix}\cos\\ \sin\end{matrix}(\phi_1+\psi_1)\,t\begin{matrix}\cos\\ \sin\end{matrix}(\phi_1-\psi_1)\,t} = 0.$$

Applying these results to equation (2·8) we have

$$\bar{\bar{\alpha}} = \tfrac{1}{2}(G_a + H_b + K_c + L_d) \qquad \text{.....(3·0).}$$

We can evaluate the constants occurring in this equation by somewhat lengthy algebra, and we find

$$\bar{\bar{\alpha}} = \frac{9\mu_1^2\mu_2^2}{2M^2r^9}\left\{\frac{13}{(\phi_1+\psi_1)^2}+\frac{1}{(\phi_1-\psi_1)^2}\right\}\cos\theta + \frac{3\mu_1^2\mu_2^2(I_1+I_2)}{8MI_1I_2r^7}\left\{\frac{9}{(\phi_1+\psi_1)^2}+\frac{1}{(\phi_1-\psi_1)^2}\right\}\cos\theta.$$

Also

$$\bar{\bar{\alpha}} = \rho\cos\theta - \tau\sin\theta,$$

where ρ and τ are the radial and transverse averaged relative accelerations respectively in the directions of r and θ increasing. Therefore

$$\left.\begin{aligned} R = M\rho &= \frac{9\mu_1^2\mu_2^2}{2Mr^9}\left\{\frac{13}{(\phi_1+\psi_1)^2}+\frac{1}{(\phi_1-\psi_1)^2}\right\} \\ &\quad + \frac{3\mu_1^2\mu_2^2(I_1+I_2)}{8I_1I_2r^7}\left\{\frac{9}{(\phi_1+\psi_1)^2}+\frac{1}{(\phi_1-\psi_1)^2}\right\} \\ T = M\tau &= 0 \end{aligned}\right\} \quad (3\cdot1),$$

where R and T are the apparent forces corresponding to the accelerations ρ and τ. Thus we see that the doublets repel each other with a force $(\lambda r^{-7} + \mu r^{-9})$, where λ and μ are positive constants which may be found by comparison with equation (3·1).

We now have to consider what is the influence of one doublet on the angular velocity of the other. Similarly to equation (3·0) we find

$$\dot{\bar{\phi}} = \dot{\bar{\phi}}_0 = \tfrac{1}{2}(G_a'' + H_b'' + K_c'' + L_d'') \qquad \text{.....(3·2).}$$

Digitized by Google

When we evaluate explicitly the constants occurring on the right-hand side of this equation we find

$$\left.\begin{aligned} G_a'' + H_b'' &= 0 \\ K_c'' + L_d'' &= 0 \end{aligned}\right\} \quad \text{.....(3·21),}$$

and so

$$\ddot{\phi} = 0 \quad \text{.....(3·22).}$$

Thus to our order of approximation the angular velocities of the two doublets remain constant throughout the interaction; and the coefficients λ and μ remain constant, their values depending on the constant values (ϕ_1 and ψ_1) of the angular velocities of the doublets. That is to say, during the whole of an encounter (provided the doublets do not approach each other too closely), the doublets appear to repel each other with a force $(\lambda r^{-7} + \mu r^{-9})$, where λ and μ are constants, given by

$$\left.\begin{aligned} \lambda &= \frac{3\mu_1^2\mu_2^2\,(I_1 + I_2)}{8I_1I_2}\left\{\frac{9}{(\phi_1+\psi_1)^2} + \frac{1}{(\phi_1-\psi_1)^2}\right\} \\ \mu &= \frac{9\mu_1^2\mu_2^2}{2M}\left\{\frac{13}{(\phi_1+\psi_1)^2} + \frac{1}{(\phi_1-\psi_1)^2}\right\} \end{aligned}\right\} \quad \text{...(3·3).}$$

§ 4. So far we have assumed that neither $(\phi_1 + \psi_1)$ nor $(\phi_1 - \psi_1)$ is small. We now consider the state of affairs when $(\phi_1 - \psi_1)$ is small. Let $\phi_1 - \psi_1 = \xi$, where ξ is small. We also write $\phi_1 + \psi_1 = \eta$. Then, since

$$\frac{1 - \cos\xi t}{\xi^2} = \tfrac{1}{2}t^2 - \tfrac{1}{24}\xi^2t^4 \text{ approx.*}$$

and

$$\frac{\xi t - \sin\xi t}{\xi^2} = \tfrac{1}{6}\xi t^3 \text{ approx.*},$$

the equations (2·6) take the form

$$\alpha - \alpha_0 = \frac{A}{\eta^2}(1 - \cos\eta t) + \frac{B}{\eta^2}(\eta t - \sin\eta t) + \tfrac{1}{2}Ct^2 + \tfrac{1}{6}D\xi t^3 - \tfrac{1}{24}C\xi^2t^4 \quad \text{.....(4·0).}$$

Corresponding to equations (2·71) we shall have

$$A = \mathrm{E}_a + \mathrm{F}_a t + G_a\cos\eta t + H_a\sin\eta t + \tfrac{1}{2}M_a t^2 + \tfrac{1}{6}N_a\xi t^3 - \tfrac{1}{24}M_a\xi^2t^4 \quad \text{.....(4·1).}$$

The new coefficients M_a, N_a, E_a, F_a, etc., may be found by substituting in (2·70) from (4·0) and comparing the result with (4·1). Thus corresponding to equation (2·8) we have

$$\begin{aligned} \ddot{\alpha} = {} & \{\mathrm{E}_a + \mathrm{F}_a t + G_a\cos\eta t + H_a\sin\eta t + \tfrac{1}{2}M_a t^2 + \tfrac{1}{6}N_a\xi t^3 \\ & \qquad - \tfrac{1}{24}M_a\xi^2t^4\}\cos\eta t \\ & + \{\mathrm{E}_b + \mathrm{F}_b t + \ldots \qquad - \tfrac{1}{24}M_b\xi^2t^4\}\sin\eta t \\ & + \{\mathrm{E}_c + \mathrm{F}_c t + \ldots \qquad - \tfrac{1}{24}M_c\xi^2t^4\}(1 - \tfrac{1}{2}\xi^2t^2) \\ & + \{\mathrm{E}_d + \mathrm{F}_d t + \ldots \qquad - \tfrac{1}{24}M_d\xi^2t^4\}\xi t \end{aligned} \quad \text{.....(4·2).}$$

* For justification of these approximations see footnote on p. 281.

Digitized by Google

We now require to take some new averages. As in § 3 we choose the time 2τ to be sufficiently large to contain a large number of periods of $\begin{matrix}\cos\\ \sin\end{matrix}\ \omega t$ and yet sufficiently small for r and θ to be effectively constant, and we then average over the time $t = -\tau$ to $t = +\tau$*. Now we are comparing the actual motion with a comparison motion in which the effects of individual rotations are smoothed out. In this comparison motion the average acceleration during the time $t = -\tau$ to $t = +\tau$ is the sum of two terms:

(i) the actual acceleration (in the comparison motion) at time $t = 0$.

(ii) a term which tends to zero as τ tends to zero.

We are trying to find the actual acceleration at time $t = 0$ in the comparison smoothed out motion. Consequently terms in our average which tend to zero as τ tends to zero are to be identified with (ii) and hence omitted. Terms which fluctuate equally positive and negative disappear as before in the smoothing out process, e.g.

$$\overline{t^2 \cos \eta t} = \frac{1}{2\tau}\int_{-\tau}^{\tau} t^2 \cos \eta t\, dt = \frac{\tau}{\eta}\sin \eta\tau + \frac{2}{\eta^2}\cos \eta\tau - \frac{2\sin \eta\tau}{\eta^3\tau}.$$

We may either say that each term in this expression fluctuates equally positive and negative and is therefore to be omitted; or we may say that $\overline{t^2 \cos \eta t} \to 0$ as $\tau \to 0$ (as is clear without integration) and hence $\overline{t^2 \cos \eta t}$ is to be omitted. Hence we write

$$\overline{t^2 \cos \eta t} = 0.$$

Similarly we put

$$\overline{t^2 \sin \eta t} = 0,$$

$$\overline{t^3 \begin{matrix}\cos\\ \sin\end{matrix}\ \eta t} = 0,$$

$$\bar{t}^n = 0, \qquad n = 1, 2, 3, 4.$$

Applying these results to equation (4·2) we have

$$\bar{\ddot{\alpha}} = \tfrac{1}{2}(G_a + H_b) + \mathrm{E}_c \qquad \text{.....(4·3).}$$

Evaluating the right-hand side of this equation explicitly we find

$$\begin{aligned}\bar{\ddot{\alpha}} = {} & \left\{\frac{117\mu_1^2\mu_2^2}{2M^2\eta^2r^9} + \frac{27\mu_1^2\mu_2^2(I_1+I_2)}{8MI_1I_2\eta^2r^7} - \frac{3\mu_1\mu_2\cos\epsilon_0}{2Mr^4}\right\}\cos\theta \\ & - \frac{27\mu_1^2\mu_2^2}{2M^2\eta^2r^9}\cos(\epsilon_0+2\chi_0)\{\cos(\epsilon_0+\theta)+\cos(\epsilon_0-\theta)\} \\ & + \frac{9\mu_1^2\mu_2^2}{2M^2\eta^2r^9}\sin(\epsilon_0+2\chi_0)\{\sin(\epsilon_0+\theta)-\sin(\epsilon_0-\theta)\} \\ & + \frac{9\mu_1^2\mu_2^2(I_1-I_2)}{8MI_1I_2\eta^2r^7}\sin(\epsilon_0+2\chi_0)\{\sin(\epsilon_0+\theta)+\sin(\epsilon_0-\theta)\} \\ & \qquad \text{.....(4·4),}\end{aligned}$$

* We have already seen that this choice of τ is possible when the doublets represent normal hydrogen atoms.

Digitized by Google

where $\epsilon_0 = \phi_0 - \psi_0$, and $\chi_0 = \psi_0 - \theta$ ...(4·41), (4·42).

Now the terms in $\begin{matrix}\cos\\ \sin\end{matrix}(\epsilon_0 + 2\chi_0)$ change sign in every quarter revolution of either doublet, the periods of revolution of the two doublets being the same. That is, these terms change sign whenever the centre ($t = 0$ in previous work) of the time over which the average is taken is made earlier or later by a quarter of the period of revolution of either doublet. Thus on the average the terms in $\begin{matrix}\cos\\ \sin\end{matrix}(\epsilon_0 + 2\chi_0)$ contribute nothing permanent to $\ddot{\bar{\alpha}}$*. Therefore we write

$$\ddot{\bar{\alpha}} = \left\{\frac{117\mu_1^2\mu_2^2}{2M^2\eta^2r^9} + \frac{27\mu_1^2\mu_2^2(I_1+I_2)}{8MI_1I_2\eta^2r^7} - \frac{3\mu_1\mu_2\cos\epsilon_0}{2Mr^4}\right\}\cos\theta.$$

Also, as before, $\ddot{\bar{\alpha}} = \rho\cos\theta - \tau\sin\theta$,

and therefore

$$\left.\begin{aligned} R = M\rho &= -\frac{3\mu_1\mu_2\cos\epsilon_0}{2r^4} + \frac{27\mu_1^2\mu_2^2(I_1+I_2)}{8I_1I_2\eta^2r^7} + \frac{117\mu_1^2\mu_2^2}{2M\eta^2r^9}\\ T = M\tau &= 0\end{aligned}\right\} \quad (4\cdot5).$$

Now similarly to equation (4·3) we have

$$\ddot{\bar{\phi}} = \ddot{\phi}_0 = \tfrac{1}{2}(G_a'' + H_b'') + E_c'' \qquad .....(4\cdot6).$$

By the first of equations (3·21) and an explicit evaluation of E_c'' this becomes

$$\begin{aligned}\ddot{\bar{\phi}} = E_c'' &= -\frac{\mu_1\mu_2}{2I_1r^3}\sin\epsilon_0 - \frac{27\mu_1^2\mu_2^2\sin\epsilon_0\cos(\epsilon_0+2\chi_0)}{4MI_1\eta^2r^8}\\ &\quad - \frac{3\mu_1^2\mu_2^2(I_1-I_2)\cos\epsilon_0\sin(\epsilon_0+2\chi_0)}{4I_1^2I_2\eta^2r^8}\\ &= -\frac{\mu_1\mu_2}{2I_1r^3}\sin\epsilon_0 \qquad .....(4\cdot7)\end{aligned}$$

by the same argument as before. Hence we cannot now say that the angular velocities of the doublets remain constant and equal to their initial values; and if $(\phi_1 - \psi_1)$ is small initially, we do not know that it will remain small.

§ 5. In order to discuss this case more fully we make the additional assumption that the two doublets have the same

* If the average value of $\begin{matrix}\cos\\ \sin\end{matrix}(\epsilon_0+2\chi_0)$ is not exactly zero, it will be very small compared with unity, and the coefficient of $r^{-n}\begin{matrix}\cos\\ \sin\end{matrix}(\epsilon_0+2\chi_0)$ in equation (4·4) is only about $\frac{1}{3}$ or $\frac{1}{4}$ of the corresponding coefficient of r^{-n} in the terms independent of $\begin{matrix}\cos\\ \sin\end{matrix}(\epsilon_0+2\chi_0)$.

Digitized by Google

moment of inertia, $I_1 = I_2 = I$ (say)*. Using Lagrange's equations of motion in ϕ and ψ, deduced from equations (2·0) and (2·1), we obtain the accurate (unaveraged) equation

$$\ddot{\epsilon} + \frac{\mu_1 \mu_2}{Ir^3} \sin \epsilon = 0,$$

where $$\epsilon = \phi - \psi.$$

Or $$\ddot{\epsilon} + n^2 \sin \epsilon = 0 \qquad \text{.....(5·0)},$$

where $$n = \sqrt{\mu_1 \mu_2 / Ir^3} \qquad \text{.....(5·01)}.$$

Thus there is a pendulum motion in ϵ, in which the length of the simple equivalent pendulum is a function of the distance apart of the doublets. Thus the motion in ϵ is analogous to the motion of a pendulum of variable length.

We now want to compare the motion in ϵ with the motion of a pendulum of slowly varying length. We assume that the length of the equivalent simple pendulum only changes by a very small fraction of itself during one period in ϵ, and that there is no resonance between the vibrations in ϵ and the rotations of the doublets. We proceed to investigate whether these assumptions are fulfilled when the doublets represent hydrogen atoms in their principal quantum orbits. If $(-e)$ and m are the charge and mass respectively of an electron, and a the radius of the principal quantum orbit in the hydrogen atom, we have $\mu = \mu' = ea$, $I = ma^2$. Therefore

$$n^2 = e^2/mr^3 \qquad \text{.....(5·1)}.$$

We find that if τ_1 is the period of vibration in ϵ for small amplitudes, and τ' the period of revolution of a doublet,

$$\tau_1 = 3{\cdot}95.10^{-4}.r^{\frac{3}{2}} \text{ secs.}, \quad \tau' = 1{\cdot}52.10^{-16} \text{ secs.} \quad (5{\cdot}20), (5{\cdot}21),$$

therefore $$\tau_1/\tau' \fallingdotseq 2{\cdot}6.x^{\frac{3}{2}}\dagger \qquad \text{.....(5·22)},$$

where $r = x.10^{-8}$, so that x is the measure of r in Ångstrom units.

Now our work is based on the assumption that r is large compared with a, i.e. x is large compared with $\frac{1}{2}$‡. Hence τ_1 is large compared with τ'. Thus the period in ϵ contains many periods of revolution of the doublets, and there is no resonance between the vibrations in ϵ and the rotations of the doublets.

Let us suppose now that the relative velocity of the doublets is such that during the time $\tau_1 \, \delta r/r$ is very small, where δr denotes

* This condition is always fulfilled when the doublets represent hydrogen atoms, for two hydrogen atoms would only have the same initial angular velocities if they had the same quantum numbers, and then they would have the same moments of inertia.

† Numerical results are calculated from formulae and data given by Sommerfeld, *Atomic Structure and Spectral Lines* (transl. Brose), and Jeans, *Dynamical Theory of Gases* (3rd ed.).

‡ $a = 0{\cdot}532$ Å. Even if r were only $10a$, τ_1 would be $30\tau'$.

Digitized by Google

the change in r during one period in ϵ. Then by equation (5·1) $\delta n/n = -\frac{3}{2}\delta r/r$, and is very small, i.e. the length of the simple equivalent pendulum changes by a very small fraction of itself during one of its own oscillations.

The conditions are now satisfied for the motion in ϵ to be analogous to the motion of a pendulum of slowly varying length and for Ehrenfest's invariant to apply. The energy integral of equation (5·0) is

$$\tfrac{1}{2}\dot{\epsilon}^2 = n^2(\cos\epsilon - \cos\gamma) \qquad \text{.....(5·3),}$$

where γ is the amplitude in ϵ*; and Ehrenfest's invariant is

$$\oint \dot{\epsilon} d\epsilon = \text{const.} \qquad \text{.....(5·4).}$$

Equations (5·1), (5·3), (5·4) can easily be shown to lead to the equation

$$r = \lambda y^{\frac{3}{2}} \qquad \text{.....(5·5),}$$

where λ is an arbitrary constant and

$$y = \sin^2\tfrac{1}{2}\gamma \int_0^{\frac{\pi}{2}} \frac{\cos^2\phi\, d\phi}{\sqrt{1 - \sin^2\tfrac{1}{2}\gamma \sin^2\phi}} \qquad \text{.....(5·51)}$$

$$= E - K\cos^2\tfrac{1}{2}\gamma \qquad \text{.....(5·52),}$$

K and E being the complete elliptic integrals of the first and second kind respectively to the modulus $\sin\frac{1}{2}\gamma$. By considering the integral in equation (5·51) we can show that r/λ or $y^{\frac{3}{2}}$ steadily increases from 0 to 1 as γ increases from 0 to π. From tables of elliptic functions we easily calculate the following values:

γ	20°	40°	80°	120°	160°
$\frac{r}{\lambda}$	0·083	0·205	0·491	0·768	0·963

Hence the curve giving r/λ as a function of γ is as shown in Fig. 2.

When we are given the value of γ corresponding to any given value of r, we can determine λ from the curve and then read off the value of γ for any value of r smaller than λ. γ and r increase and decrease together. If γ is real for any given value of r, γ is real for any smaller value of r.

Now before two rotating doublets with the same moment of inertia enter each other's fields their angular velocities will be the same if they have the same quantum number. That is, when r is

* The case of one doublet gaining complete revolutions on the other, corresponding to the equivalent pendulum swinging right round, would be represented in the analysis by γ imaginary or $|\cos\gamma| > 1$.

Digitized by Google

infinite, $\dot{\epsilon} = 0$. Now γ is the value of $|\epsilon|^*$ when $\dot{\epsilon} = 0$. Therefore, when r is infinite, γ is the angle between the doublets and is therefore real. Therefore r/λ is finite. Therefore λ is infinite. Therefore when r is finite, $r/\lambda = 0$, $\gamma = 0$, and $\epsilon = 0$. This conclusion is only true if the preceding work is valid for all values of r greater than some definite finite value. We now investigate the limits of validity of our work.

When the doublets correspond to hydrogen atoms in their principal quantum orbits, we must have $r > r_1$, where r_1 is moderately large compared with a, the radius of a hydrogen atom.

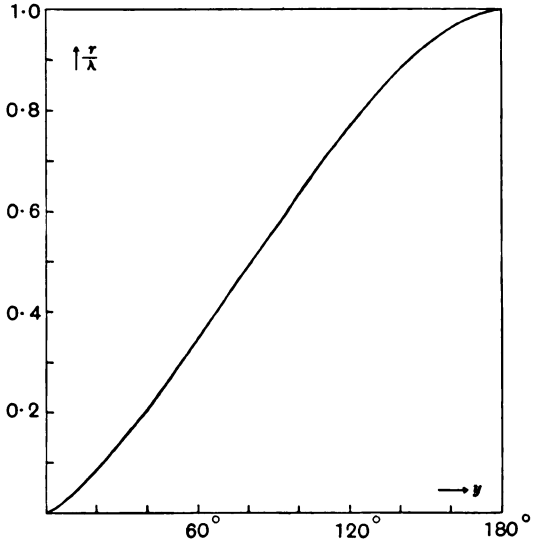


Fig. 2

Suppose for definiteness we take $r_1 = 10a = 5$ Å. This satisfies Debye's condition mentioned in § 1 for considering the field of a hydrogen atom as due to a rotating doublet. The force on an electron due to an electron or proton at a distance of $10a$ is only $\frac{1}{100}$ of the force on it due to its own proton. Hence the "coupling" is small. The distance travelled in one revolution of either doublet, when the doublets have the kinetic theory velocity appropriate to hydrogen atoms at 0° C. is $3{\cdot}8.10^{-11}$ cm., or less than $\frac{1}{1000}$ of their distance apart. Thus the condition mentioned just before equation (2·6) is satisfied. Also a choice of τ is possible for averaging purposes; and the condition that τ_1/τ' is large is satisfied.

Also the condition that $\delta r/r$ is to be small leads to the condition

* I.e. the absolute value of ϵ, when ϵ is taken to satisfy $-\pi < \epsilon \leqslant \pi$.

Digitized by Google

$r < R_1$, where R_1 is small compared with 10,000 Å., $R_1 = 500$ Å. (say). Then our treatment of the motion is valid when $r_1 < r < R_1$.

Now we can see from Fig. 2 that whatever value γ had when $r = R_1$, when $r < \frac{1}{10}R_1$, $\gamma <$ about 24°, i.e. $1 \geqslant \cos\gamma \geqslant 0{\cdot}9$. Now $\cos\epsilon$ varies between $\cos\gamma$ and 1, hence very little error will result from putting $\cos\epsilon = 1$ for that part of the motion for which $r < \frac{1}{10}R_1$.

Thus, if $r_1 < r < \frac{1}{10}R_1$, or 5 Å. $< r <$ 50 Å., the two rotating doublets move as if there were a force of repulsion (R) between them given by

$$R = -\lambda_4 r^{-4} + \lambda_7 r^{-7} + \lambda_9 r^{-9} \qquad \text{.....(5·6),}$$

where

$$\lambda_4 = \tfrac{3}{2}\mu_1\mu_2, \quad \lambda_7 = 27\mu_1^2\mu_2^2/4I\eta^2, \quad \lambda_9 = 117\mu_1^2\mu_2^2/2M\eta^2$$
$$\text{.....(5·61), (5·62), (5·63),}$$

obtained by putting $I_1 = I_2 = I$ and $\cos\epsilon = 1$ in equation (4·5).

From equation (4·7) and the similar one for $\bar{\psi}$ we see that when $I_1 = I_2$, $\ddot{\phi} + \ddot{\bar{\psi}} = 0$. Also the relative angular velocity is so small that one doublet gains an angle γ ($< \pi$) on the other only in the course of a large number of revolutions of either (since τ_1/τ' is large)*. Hence the angular velocities of the two doublets remain approximately constant and equal to their initial values, which are, of course, equal. Hence the λ's in equation (5·6) are constants, their numerical values being $\lambda_4 = 9{\cdot}68.10^{-36}$, $\lambda_7 = 1{\cdot}62.10^{-60}$, $\lambda_9 = 4{\cdot}34.10^{-79}$. We notice that when $r = 1$ Å. ($= 10^{-8}$ cm.) the repulsive term varying as r^{-9} is very small compared with the repulsive term varying as r^{-7} (the ratio being about $\frac{1}{3}\frac{1}{7}\frac{1}{0}$); the term in r^{-9} diminishes more rapidly with increasing distance than the term in r^{-7}. Thus at the distances for which our work is valid, we may write without sensible error

$$R = -9{\cdot}68.10^{-36}r^{-4} + 1{\cdot}62.10^{-60}r^{-7} \qquad \text{.....(5·7).}$$

This is a law of force of the same form as that obtained empirically by Lennard-Jones for helium, argon, and neon.

The attractive force in R just balances the repulsive force (i.e. $R = 0$) if $r \fallingdotseq 0{\cdot}55$ Å. This distance is much less than the minimum distance for which our work is valid. All we can deduce is that at distances for which our work is valid the attractive force preponderates. This excessive attractive force is not in accordance with the empirical results of Lennard-Jones. There are two possibilities, it may be due to the specialization of the mathematical

* This statement justifies the expansions of $\begin{matrix}\cos\\ \sin\end{matrix}\xi t$ in § 4 in the case when the rotating doublets represent normal Bohr hydrogen atoms, for we only require these expansions to be true for moderately few periods of revolution of each atom including the time $t = 0$; hence ξt is small compared with γ, itself less than π.

Digitized by Google

model; or, as a gas of hydrogen atoms is unstable and forms into a gas of ordinary hydrogen, it may be that the attractive force is greater in the case of hydrogen than in the case of the three monatomic gases helium, argon, and neon.

A useful way of comparing repulsive fields is by means of "diameters" as defined by Lennard-Jones*. If σ is the "diameter" corresponding to a repulsive force $\lambda_n r^{-n}$, we have

$$\sigma = \{\lambda_n/2{\cdot}06.10^{-16}\,(n-1)\}^{1/(n-1)}.$$

Applying this conception to the repulsive term in equation (5·7), we find $\sigma = 3{\cdot}31$ Å. This is of the same order of magnitude as the values obtained empirically by Lennard-Jones for helium ($\sigma = 3{\cdot}124$ Å.), neon ($\sigma = 4{\cdot}300$ Å.), and argon ($\sigma = 7{\cdot}053$ Å.). The value of n ($n = 7$) is rather smaller than those obtained empirically by Lennard-Jones for helium ($n = 14$), neon ($n = 11$), and argon ($n = 9$). It is to be noticed that the hydrogen atom is very much more unsymmetrical than the atoms of the three inert gases mentioned; on this account we should expect the index n to have a smaller value than in the case of these three inert gases. The agreement between the calculated and observed repulsive forces is quite good in view of the model chosen.

§ 6. *Summary.* We may summarize our conclusions as follows. If the rotating doublets have quite different angular velocities initially, then they repel each other with a force (R) given by

$$R = \lambda r^{-7} + \mu r^{-9} \qquad \text{.....(6·0),}$$

where
$$\lambda = \frac{3\mu_1^2\mu_2^2\,(I_1+I_2)}{8I_1I_2}\left\{\frac{9}{(\phi_1+\psi_1)^2}+\frac{1}{(\phi_1-\psi_1)^2}\right\} \qquad \text{...(6·01),}$$

and
$$\mu = \frac{9\mu_1^2\mu_2^2}{M}\left\{\frac{13}{(\phi_1+\psi_1)^2}+\frac{1}{(\phi_1-\psi_1)^2}\right\} \qquad \text{.....(6·02),}$$

ϕ_1 and ψ_1 being the (constant) angular velocities of the two doublets.

If the doublets have the same angular velocities initially, and the same moments of inertia, then over a certain range of r we have

$$R = -\lambda_4 r^{-4} + \lambda_7 r^{-7} + \lambda_9 r^{-9} \qquad \text{.....(6·1),}$$

where

$$\lambda_4 = \tfrac{3}{2}\mu_1\mu_2, \quad \lambda_7 = 27\mu_1^2\mu_2^2/16I\omega^2, \quad \lambda_9 = 117\mu_1^2\mu_2^2/8M\omega^2$$
$$\text{.....(6·11), (6·12), (6·13),}$$

and ω is the common value of the angular velocities of the doublets. When the doublets correspond to hydrogen atoms in their principal

* J. E. Jones, *Proc. Roy. Soc.* A, vol. CVI, pp. 452 *et seq.* (1924).

Digitized by Google

quantum orbits, the range of distance becomes 5 Å. to 50 Å. and the formula for R reduces to

$$R = -9{\cdot}68.10^{-36}r^{-6} + 1{\cdot}62.10^{-60}r^{-7} \quad \ldots\ldots(6{\cdot}2).$$

This is a law of force of the type found empirically by Lennard-Jones for helium, neon, and argon. The attractive term in this formula is larger than the attractive terms found by Lennard-Jones. The repulsive term, however, which leads to a "diameter" of 3·31 Å., is in very satisfactory agreement with the repulsive terms found by Lennard-Jones.

Digitized by Google

A Theory of the Action of X-rays on living Cells. By J. A. CROWTHER, Sc.D., St John's College.

[*Received* 13 March, *read* 15 March, 1926.]

The absorption by matter of energy from a beam of X-rays follows laws which are now well known. The first stage is the ejection from the absorbing atom of a high speed electron. This electron, in turn, produces pairs of ions from some of the molecules through which it passes until its energy is all spent. The process is essentially a discontinuous one in space, and the proportion of atoms affected at a given time is always exceedingly minute, even with an intense beam of radiation. With a beam of average intensity an individual atom would suffer ionisation, on an average, about once in a million years.

This discontinuous nature of the absorption of X-rays requires to be taken into account in considering the effect of X-rays on living matter. I have already pointed out* that some of the important structures in a living cell are so minute that the probability of their being affected by the radiation in a given short interval of time, say one minute, is quite small. In the particular instance discussed in the paper referred to, this probability amounted to about 0·058 per minute, on the assumption that ionisation in the particle was sufficient to bring about a biological change. If, on the other hand, we assumed that the actual expulsion of a high-speed electron from the particle was necessary to bring about a biological action, the probability would be even smaller. There does not appear to be sufficient evidence to allow of discrimination between these two possibilities. Such discrimination is not necessary for the purpose of the present paper.

Let λ be the probability that a particle of a given kind will be "hit" when unit quantity of X radiation passes through the particles, and let N be the number of such particles present initially. The total number of particles y_0 which are still unaffected by the radiation after a total dose q is given by

$$y_0 = N\epsilon^{-\lambda q}.$$

If a single hit is sufficient to put the particle out of action this equation will give the number surviving a dose q of X radiation. It was suggested in a former paper† that an exponential curve of this kind represented fairly adequately the results of experiments of Strangeways and Oakley‡ on the action of X-rays on tissue cells.

* J. A. Crowther, *Proc. Roy. Soc.* B, vol. XCVI, p. 207 (1924). † *Loc. cit.*
‡ *Proc. Roy. Soc.* B, vol. XCV, p. 373 (1923).

Digitized by Google

The changes studied by Strangeways and Oakley in these experiments were produced by comparatively small doses of X-rays. Many biological changes require much heavier doses of radiation and the size of particle required for their explanation on the above assumption would be much smaller than we have any evidence for assuming to exist. Moreover, the relation between survival and dosage is certainly not always exponential. More usually, no apparent effect is produced unless the dose exceeds some critical value: the number of survivors then decreases rapidly with increasing dosage, though a few stray specimens survive much larger doses.

A fairly obvious suggestion is that, in such cases, a succession of hits is necessary to produce the effect which is being studied. Many of the cell structures are known to be electrically charged. It may, perhaps, be necessary to neutralise either the whole, or a definite fraction of this charge to produce a change in the structure, just as the neutralisation of the charge on a colloid brings about precipitation. The object of the present paper is to follow out the results of this suggestion.

Since the particles we are considering are large compared with an atom we can assume that the probability of a second, or third hit is the same as that for the first.

Let λ be the probability that a particle of the given kind will be hit once by a quantity of X-rays measured by q, and let y_n be the number which have already been hit n times. The number of these which will be hit again by a small additional dose dq will be $\lambda y_n . dq$. These will pass into the state $(n+1)$. With the same dose dq, however, a number $\lambda y_{n-1} . dq$ of the particles which have already been hit $n-1$ times will be hit again and so pass into the state n. Thus we have

$$dy_n = \lambda y_{n-1} . dq - \lambda y_n . dq,$$

$$\frac{dy_n}{dq} = \lambda y_{n-1} - \lambda y_n,$$

$$(D + \lambda)\, y_n = \lambda y_{n-1},$$

writing D for d/dq. Also $y_0 = N\epsilon^{-\lambda q}$. Whence we have

$$y_n = N \frac{\lambda^n q^n}{n!} \epsilon^{-\lambda q}.$$

If n successive hits are required to put the particle out of action, the number surviving a dose q is thus the sum of all the particles in the states for which the number of hits is less than n, i.e.

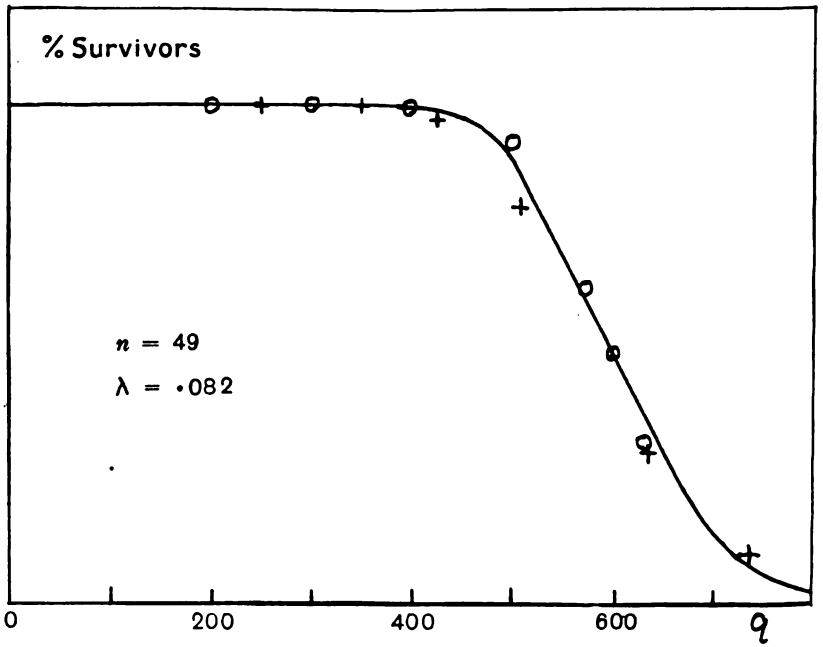
$$\begin{aligned} Y &= y_0 + y_1 + \ldots + y_{n-1} \\ &= N\epsilon^{-\lambda q}\left(1 + \lambda q + \frac{\lambda^2 q^2}{2!} + \ldots + \frac{\lambda^{n-1} q^{n-1}}{n-1!}\right). \end{aligned}$$

Digitized by Google

The quantity within the bracket consists of the first n terms in the expansion of $\epsilon^{\lambda q}$ and can be put in the form

$$\epsilon^{x}\left(1-\frac{\Gamma_{x}(n)}{\Gamma(n)}\right),$$

where $x = \lambda q$, $\Gamma(n)$ is the complete gamma function, and $\Gamma_x(n)$ the incomplete gamma function. The ratio $\Gamma_x(n)/\Gamma(n)$ has recently been tabulated by Professor Karl Pearson for an extensive range of values. The computation of the survival curve for given values of n and λ thus presents no difficulties. It is plotted in the accompanying figure, for the values $n = 49$, $\lambda = 0{\cdot}082$.



A curve of this general shape represents the results of several researches into the action of X-rays on living cells. Unfortunately the numerical results are not sufficiently precise to enable the agreement to be tested numerically. It was thought that suitable material for a test might be found among the protozoa. I have accordingly been making observations on the effect of X-rays on a pure culture of *Colpidium Colpoda*, very kindly supplied to me by Dr Ward Cutler, of Rothamsted. The results of two complete series of such experiments are indicated by the circles and crosses in the figure, the ordinates representing the percentage of individuals surviving a dose of radiation measured by the corresponding

Digitized by Google

abscissae. The experiments are tedious, and somewhat difficult. The quantity of radiation is measured by ionisation methods, and individual readings of the quantity have a probable error of some 2 or 3 %. Within these limits the two series of experiments agree with each other, and with the theoretical curve with reasonable accuracy. Further experiments are being made with the object of fixing the position of the experimental points with greater precision. Assuming, for the moment, that the theoretical curve drawn is the best fit for the experimental results, they would indicate that Colpidium requires a succession of 49 hits on some particular structure before it is killed by the X radiation. Protozoa are, in fact, highly resistant to the rays, and, as far as I am aware, this is the first time that a lethal effect produced by X-rays has been observed in protozoa of this type.

The experiments have been made possible by a grant from the Government Grant Committee of the Royal Society, to whom I tender my best thanks. I am also deeply indebted to Mr Budden who has very kindly prepared the cultures used in the experiments.

Digitized by Google

On the Attenuation of Wireless Signals in Short Distance Overland Transmission. By J. A. RATCLIFFE, B.A. and M. A. F. BARNETT. (Communicated by Mr E. V. APPLETON.)

[*Read* 3 May 1926.]

1. *Introduction.*

Recent work on the explanation of fading and allied phenomena* has directed attention to the fact that the disturbance constituting a wireless signal travels, in general, from the transmitter to the receiver by two different paths. One of these disturbances passes over the curved surface of the earth and has been called the "ground ray" or the "direct ray." The other passes up into the higher layers of the atmosphere, is there deviated by the ionised "Heaviside layer," and is finally returned to the ground at the receiving station. This ray has been called the "indirect" or "atmospheric" ray. It is also shown (*loc. cit.*) that close to a transmitter, any energy which travels by way of the atmospheric ray is negligible in comparison with the energy travelling by the ground ray.

As the distance from the transmitter is increased, the ground ray is reduced in strength more rapidly than the atmospheric ray, and finally, at great distances, the atmospheric ray alone is of importance. The distance at which the atmospheric ray first begins to be comparable in strength with the ground ray varies with the wave length. It has been shown (*loc. cit.*) how, from "fading" measurements, it is possible to determine the relative effects of the two rays at any given point.

In this paper we shall consider in greater detail what happens to the ground ray as it passes from transmitter to receiver.

2. *Attenuation of the Ground Ray.*

If a sending antenna of effective height h_s is placed in free space, or on the surface of a perfect conductor, then the value of oscillating E.M.F. produced in a receiving antenna of effective height h_r at a given point in the equatorial plane of the transmitter, and at a distance d from it, is given by

$$E = k\frac{h_s h_r i}{\lambda d},$$

* Appleton, *Proc. Phys. Soc.* vol. XXXVII, Part 2 (Feb. 1925); Appleton and Barnett, *Nature* (March 7, 1925); *Electrician*, p. 398 (April 3, 1925); *Proc. Camb. Phil. Soc.* vol. XXII, Part 5, p. 674; *Proc. Roy. Soc.* A, vol. CIX, p. 621.

Digitized by Google

where k is a constant, i is the current in the sending antenna and λ is the wave length. In this expression it is assumed that h_s and h_r are both small compared with the wave length.

For such a system the received E.M.F. falls off inversely as the distance, and the product of E.M.F. and distance is a constant.

In the actual case of transmission over the earth's surface, both the transmitter and receiver are on the surface of a sphere. The waves constituting the ground ray can only travel from the one to the other by diffractive bending round the surface of this sphere. At the receiver the strength of the signal will therefore be less, in general, than it would be if the earth were plane. To pass from the expression for transmission over a plane conductor to that for transmission round a spherical conductor, whose radius is equal to that of the earth, it is necessary to multiply by a factor $\phi(\lambda, d, R)$ which is a function of the wave length, the distance (d) between the transmitter and receiver, and the radius (R) of the earth. This factor has been calculated from theoretical considerations by several mathematicians*.

Although the results obtained by them differ from one another, yet it is clear that the correction due to this factor is negligible, provided the transmission takes place over distances so short that the tangent plane through the transmitter is not distant more than one wave length from the surface of the sphere at the receiver. If, therefore, measurements are made sufficiently near the transmitter, this factor may be neglected. If the measurements are made at greater distances, the factor differs appreciably from unity, but provided the distances are not too great, the correction is still small and can be made from the theory sufficiently accurately.

So far we have considered only the case of a perfectly conducting sphere. For an imperfect conductor, the wave surface near the boundary is distorted, so that the diffractive bending factor $\phi(\lambda, d, R)$ must involve the electrical constants of the sphere, ρ the specific resistance, and k, the dielectric constant. This factor will then become $F(\lambda, d, \rho, k, R)$.

For an imperfect conductor there is a further source of attenuation due to dissipation of energy by the currents flowing in the conductor. The attenuation of the waves due to this cause will depend on the specific resistance of the ground, on its dielectric constant, on the wave length of the wave, and on the distance over which it has passed. It may be represented by a factor $f(\rho, k, \lambda, d)$. Thus, finally, the electric field due to an antenna situated on the

* Love, *Phil. Trans.* A, vol. ccxv, p. 105 (1905), a complete bibliography up to 1915 will be found in this paper; G. N. Watson, *Proc. Roy. Soc.* A, vol. xcv, p. 83 (1918) and p. 546 (1919); Van der Pol, *Phil. Mag.* vol. xxxviii, p. 365 (1919).

Digitized by Google

surface of the earth is given, at a distance d, by an expression of the form

$$E = k \frac{h_s h_r i}{d\lambda} F(\lambda, d, \rho, k, R) f(\rho, k, \lambda, d).$$

The product of E and d is no longer constant, as in the case of a plane conducting earth, but is given as a function of the distance, by the expression $F(\lambda, d, \rho, k, R) f(\rho, k, \lambda, d)$. We may refer to this as the "attenuation factor" of the waves. The part represented by $F(\lambda, d, \rho, k, R)$ may be called the "diffractive attenuation" and the part $f(\rho, k, \lambda, d)$ the "flat ground attenuation."

3. *Measurement of the "Flat Ground Attenuation."*

The electric intensity of a wireless signal at a given point is, in general, a complicated function of the distance from the transmitter. It depends on the strength of the atmospheric ray and on the strength of the ground ray, the latter being given as we have seen by $k \frac{h_s h_r i}{d} Ff$. By measurements on such a signal we can only obtain data which depend (a) on the strength of the atmospheric ray, (b) on the "diffractive attenuation," and (c) on the "flat ground attenuation." The measurements cannot be interpreted so as to give the individual contribution of these three effects. If two of the effects can be eliminated, then experiments will lead directly to a knowledge of the third. Now it has been shown that if measurements are confined to points near the transmitter, then (a) the effect of the atmospheric ray, and (b) the effect of "diffractive attenuation," are both negligible. Such measurements will therefore give directly the effect of "flat ground attenuation," i.e. they will give the factor $f(\rho, k, \lambda, d)$. The measurements described in this paper were made on waves which had travelled over distances which were "short" in the above sense. From them it is possible to determine directly the "flat ground attenuation."

4. *Previous Measurements of Attenuation.*

It is noteworthy that there have been very few determinations of the attenuation of wireless signals. This is probably due to the fact that it is only comparatively recently that methods for quantitative measurement of weak signals have been available.

Those measurements which have been made may be divided into two classes:

(a) short distance measurements,

and (b) long distance measurements.

The second type of measurement has been made much more

Digitized by Google

frequently, as is quite natural, for, of course, these are the measurements which are of commercial interest. As has been shown above, it is unfortunately not possible to deduce the individual contribution of the various attenuating factors from these measurements. So far as the authors are aware, only two series of measurements over short distances have been made. Taylor and Duddell*, in 1905, measured the attenuation of signals for distances of a mile or two over land and of 60 miles over sea-water. Bown and Gillette†, in 1924, measured the attenuation over land for distances up to 70 miles. The latter measurements are complicated by the fact that they were made over a mountainous area. The measurements of Taylor and Duddell for attenuation over land do not extend far enough for the ground attenuation to be estimated accurately. It is obvious that more data of this nature are urgently needed, in order to study the ground attenuation of wireless waves.

5. *Deductions from the Measurements.*

(*a*) *Experimental test of Sommerfeld's attenuation formula, and determination of the electrical constants of the ground.* Theoretical expressions for the attenuation due to the finite conductivity of the ground $f(\rho, k, \lambda, d)$ have been made by Zenneck‡ and Sommerfeld§.

Zenneck treats the case of plane waves and his treatment is not valid for points near the transmitter, such as will be considered here.

Sommerfeld's treatment applies to the case of waves spreading from a point source, and gives their "flat ground attenuation" in terms of the specific resistance and dielectric constant of the ground. It is thus possible to compare the observed form of the function $f(\rho, k, \lambda, d)$ with that obtained theoretically by Sommerfeld. We can make the theoretical curve fit the experimental curve by a suitable choice of the constants ρ and k. If this is done for two different wave lengths, and the values obtained for the constants are the same for both, then this lends strong support to the theory. If the values obtained for ρ and k also agree with the values found by other methods, then the support is strengthened, and the measurements constitute a verification of Sommerfeld's theory.

It so happens that, if the wave length is suitably chosen, either the resistivity, or the dielectric constant, becomes unimportant and can be neglected in the formula. Then from an observed attenuation

* *Journ. Inst. of Electr. Eng.* vol. XXXV, p. 321 (1905); *Electrician*, vol. LV, p. 409 (1905).

† *Proc. Inst. Rad. Eng.* (Jan. 16, 1924).

‡ *Ann. der Phys.* vol. XXIII, p. 846 (1907).

§ *Ann. der Phys.* vol. XXVIII, p. 665 (1909); *Jahrbuch der drahtlosen Telegraphie und Telephonie*, vol. IV, p. 157 (1910).

Digitized by Google

curve the remaining constant can be found. By working with waves of medium length (about 1000 metres) the resistivity of the ground can be found, while by working with very short waves the dielectric constant is arrived at. Van der Pol* has pointed out that this can be seen simply in the following manner. The currents which flow in the ground and are responsible for the attenuation of the wave, are of two kinds, conduction currents and displacement currents. The conduction current is proportional to $\frac{E}{\rho}$, where E is the electric force in the wave and ρ is the resistivity of the ground. The displacement current is proportional to $\frac{k}{4\pi}\frac{dE}{dt}$, k being the dielectric constant of the ground.

By assuming approximate values for ρ and k the relative importance of these two currents can be found. Such values may be taken as

$$\rho = 10^{-8} \text{ E.S.U.}\dagger,$$

$$k = 10 \text{ E.S.U.}\ddagger.$$

Then the conduction current is proportional to $E/10^{-8}$, and if E is sinusoidal with angular frequency $p = 2\pi n$, the displacement current is $\frac{10}{2}En$. Thus if n is of the order $\frac{2}{10} \times 10^8$ the attenuation is due to both currents equally, while if n is much less than this "critical" frequency the attenuation is mainly due to the conduction current. The "critical" frequency corresponds to a wave length of about 15 metres and for wave lengths greater than this the conduction current is the determining factor, while for shorter waves the displacement current is the more important. It may here be noted that Van der Pol uses a different value for ρ and so arrives at a different critical frequency. We see, therefore, that it is to be expected that measurements on signals of ordinary wave lengths (above 300 metres) will give information about the resistivity of the ground. It is impossible to make measurements on wave lengths very much shorter than 15 metres, so as to arrive at the dielectric constant of the ground directly, and so it is necessary to make measurements on wave lengths near 15 metres where both dielectric constant and resistivity are important. Then by substituting the value for the resistivity found by long wave measurements, the dielectric constant can be deduced.

(*b*) *Variation of attenuation with wave length.* Sommerfeld's formula for the electrical intensity produced by a transmitter is very complicated and must be computed for each wave length

* *De Ingenieur*, No. 38 (1925).

† Smith-Rose and Barfield, *Experimental Wireless*, p. 737 (Sept. 1925); *Proc. Roy. Soc.* A, vol. CVII, p. 587.

‡ Zenneck, *Ann. der Phys.* vol. XXIII, p. 859 (1907).

Digitized by Google

separately. It does not yield a simple relation between wave length and attenuation. The simplest formula which relates attenuation to wave length is one deduced semi-empirically by Austin* for long distance transmission over sea. Modifications of this formula have been given by Fuller† and by the Marconi Co.‡ which are said to represent the facts more accurately. All these formulae, however, are semi-empirical in nature, and apply to long distance over-sea transmission. It is not to be expected that they will hold for short distance over-land transmission, but they have often been assumed as first approximations, and it is interesting to see how nearly they express the facts, especially as regards change of attenuation with wave length.

(*c*) *Reflection coefficient of the Heaviside layer.* Professor Appleton and one of the authors have pointed out§ that if the attenuation of the ground ray is known, then from fading experiments it is possible to calculate the percentage of the atmospheric ray which is returned from the upper layer. This percentage they call the reflection coefficient of the Heaviside layer. By calculations which are given in detail in their paper they deduced the reflection coefficient by using the most reliable results then available for the attenuation of the ground ray. By applying the results of the present paper it is now possible to obtain a more accurate value for this reflection coefficient.

6. *Experimental Details.*

Measurements have been made to determine the "flat ground attenuation" of waves of lengths 1600 m. and 360 m. and measurements are in progress on waves of length 7 m.

The experimental arrangements for the various wave lengths are entirely different and will be described separately.

(*a*) 1600 *metres.* The signal strength of the Daventry station of the B.B.C. (5 XX) was measured at increasing distances up to 70 miles from the transmitter. The places chosen were on the line running directly eastward from Daventry. The ground consists of clays and oolitic limestones and ironstones up to a distance of about 35 miles, and of chalk for the greater distances. Positions for observations were always chosen to be on open ground, and away from telephone wires (wherever possible more than one wave length away).

The signals were received on a loop aerial, were amplified with a constant high frequency amplifier, and were then measured as a

* *Bull. of the Bureau of Standards*, vol. VII, No. 3.
† *Electrician*, p. 154 (May 7, 1915).
‡ *Journ. Inst. Electr. Eng.* vol. LXIII, p. 977 (1925).
§ *Proc. Roy. Soc.* A, vol. CIX, p. 367.

Digitized by Google

high frequency E.M.F. by means of a Moullin voltmeter (*M*, Fig. 1).

The loop consisted of a very large single turn loop measuring 20 m. × 5 m. supported on poles. It was always directed by compass bearings towards the Daventry station. The amplifying set was placed a foot or two from the ground, at the centre of the loop. Tests with the loop perpendicular to the direction of propagation of the waves showed that the direct "antenna effect" of the loop was negligible. The loop was tuned to 1600 m. by means of an inductance in series with it and a condenser across the combination.

The potential difference established across the condenser was applied between the grid and filament of a valve (Fig. 1). This valve was supplied with negative grid bias of such an amount that

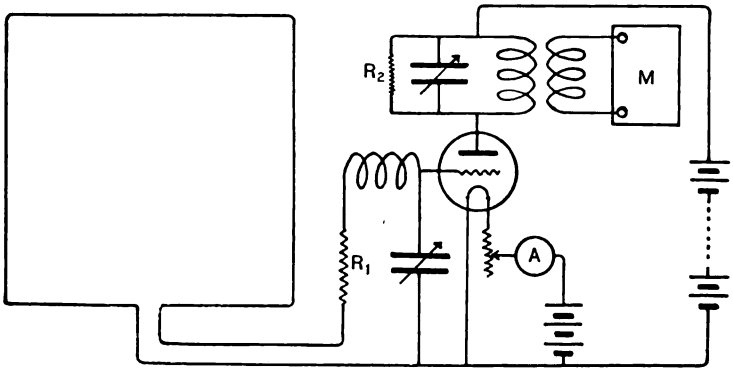


Fig. 1. Amplifier for 1600 metres.

the strongest signals received did not cause the grid to become positive with respect to the filament. The damping introduced across the loop was therefore independent of the signal amplitude. The value of the high tension E.M.F. applied to the anode was adjusted so that the valve was working on a straight part of its characteristic. This ensures that the amplification is linear and the output is proportional to the input. The filament current was always adjusted to the same value by means of a rheostat and ammeter. The value chosen was that giving maximum amplification, so that a first order small variation in filament current should make only a second order variation in amplification.

The output of the valve was passed through a tuned transformer to the Moullin voltmeter. Owing to interelectrode capacities, the valve would have produced oscillations when the two tuned circuits were syntonised. To prevent this, resistances R_1 and R_2 were put in both circuits, sufficient to render the valve quite

Digitized by Google

stable. This had the additional advantage of flattening the tuning of the circuits.

The scale on the Moullin voltmeter only covered the range from 0·5 to 1·5 volts. Owing to this small available ratio it was necessary to compare the signal strengths at the different places in pairs, the amplification being changed between one pair and the next. This change of amplification was produced by varying the resistances R_1 and R_2 in the oscillatory circuits.

It is of course essential, in this method of measurement, that the overall amplification of the apparatus should be the same at the two places to be compared. The above precautions were taken to ensure this. The extent to which the amplification did actually vary may be seen from the following figures. Measurements were made at a certain place at 11 a.m.; the mean readings on the voltmeter were 1·01 and 0·73 corresponding to two different values of the resistances R_1 and R_2. The apparatus was then taken about 40 miles and readings were taken at two other places. At 4 p.m. observations were again made at the first place and the values obtained were 1·06 and 0·76. This is typical and shows that during one day's observations the apparatus remains constant to within about 5 per cent. To compare observations on different days we may quote the following. Values were obtained for the ratio of signal strength at two places on four different occasions. Between some of the observations changes were made in the apparatus. The four values for the ratio were 1·56, 1·44, 1·64, 1·61. The deviation from the mean is here about 7 per cent. These results include possible variations in the actual strength of the signal transmitted from Daventry, and possible variations in the state of the ground. They are taken as showing that the apparatus was reasonably constant and that the precautions taken were effective.

For measurements within about 7 miles of Daventry it was possible to use a smaller loop wound with several turns and to dispense with the amplifier, the oscillating voltage being read directly on the Moullin voltmeter. Both this direct method and the amplification method were used for the distances less than 7 miles and the sort of agreement obtained can be seen from the points plotted in Fig. 4.

The possible variation in the strength of the signal transmitted is known from the readings of aerial current which are regularly taken at the transmitting station. We are informed by the B.B.C. that the current is constant to within one per cent.

The Daventry station transmits telephony on the modulated carrier wave system. The modulation was found to cause rapid variations of about 3 per cent. in the received voltage. The mean value of the voltage over about 30 seconds was taken as the value corresponding to absence of modulation.

Digitized by Google

(*b*) 360 *metres*. The telephony signals from the London station of the B.B.C. (2 LO) were measured at places near the line running North from London, up to a distance of about 70 miles. The soil consists of London clays for the first 20 miles and of chalk beyond this. The precautions with regard to telephone wires, which were mentioned above, were also observed in this case.

The signals were received on a small loop aerial and were amplified and detected so as to give a deflection on a galvanometer scale. The loop was then turned at right angles to the direction of propagation of the waves so that the London signal was cut out, and a signal from a calibrated oscillator was fed into the amplifier and its strength was varied until it produced the same galvanometer deflection as the London signal. The relative strength of the signal was then known from the reading of the calibrated oscillator.

The loop was about 1 metre square and was tuned to the wave length of 2 LO. A resistance of 100,000 ohms was shunted across the tuning condenser to decrease the sharpness of tuning. A heterodyning oscillation was induced into the aerial circuit from a separate oscillator and this, together with the carrier wave of the incoming signal, was passed through a triode employing grid condenser detection. The resulting heterodyne whistle was amplified by a two stage audio-frequency amplifier and was finally rectified by a crystal. The rectified current was passed through a galvanometer and gave a measure of the strength of the incoming signal. The heterodyning oscillator and its batteries were enclosed in tin boxes, and this shielded them sufficiently to prevent any appreciable direct pick-up of the oscillation by the loop. This is essential if the strength of the heterodyne, and so the amplification, is to remain constant for all positions of the loop. It was also necessary to shield the amplifier itself by putting it, and its batteries, in a box lined with galvanised iron. It was then found that the amplification was independent of the position of the loop, within the desired order of accuracy.

It was necessary to have a very weak signal of variable strength, and whose relative intensity could be measured, as the comparison signal, to be fed into the apparatus and compared with that received from 2 LO. This was produced by an oscillator using toroidal inductances as described by Dye* which was enclosed with its batteries in a tin box. The oscillatory circuit was coupled to the aerial circuit by means of an intermediary circuit. The coupling at the oscillatory circuit end was by means of a single turn of wire, and at the aerial circuit end was by a transformer, the primary and secondary coils of which each consisted of one turn. The current in the intermediate circuit, and hence the E.M.F. induced in the aerial circuit, was varied by means of a variable

* *Proc. Inst. Electr. Eng.* vol. LXIII, p. 597 (1925).

Digitized by Google

resistance R (Fig. 2) in the intermediate circuit. The variation of resistance was not fine enough to give accurate settings and so four or five fixed values of resistance were chosen such that the induced signal strength for each value could be made to overlap that for neighbouring values by varying the filament current in the oscillating triode. By this combination of variable filament current with four or five fixed values of resistance in the intermediate circuit, it was possible to vary the induced E.M.F. over the required range. It is essential that the coupling between the local-signal oscillator and the amplifier should always be the same, and should be independent of the position of the loop. This is the case provided there is no direct pick-up by the loop. It was found possible to

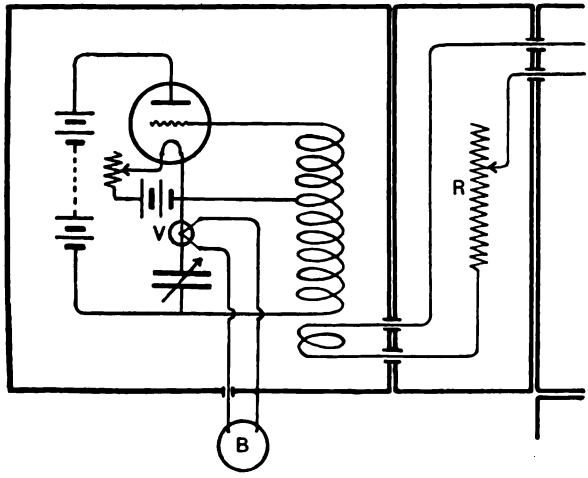


Fig. 2. Local signal generator.

prevent this direct pick-up by using the toroidal coils mentioned above, and by enclosing the oscillator and batteries in a metal box. The direct pick-up was then negligible in comparison even with the weakest induced signals used, and the induced E.M.F. was independent of the position of the loop.

The current in the oscillating circuit was measured by means of a vacuum thermojunction (V) (Fig. 2) placed directly in the circuit. Leads were brought from the junction to a galvanometer (B) outside the shielding box. The direct current calibration of the junction was used, since tests with a Moullin voltmeter indicate that this is still valid at frequencies corresponding to 360 m. wave length.

The E.M.F. induced in the aerial circuit is proportional to the current in the oscillatory circuit for a fixed resistance R and so its

Digitized by Google

relative strength is known from the thermojunction readings. The effect of using the different values of resistance R is found by comparing them together in pairs. This is done by finding what current, as recorded by the thermojunction, is needed with one of them, to give the same aerial E.M.F. as a certain current in conjunction with the next resistance. In this way a multiplying factor is found which reduces currents used in conjunction with any one resistance, to currents used with the first resistance.

The process of taking a reading is as follows. The direction of the London station is found by means of the position of the loop for minimum intensity. The loop is then turned to the position of

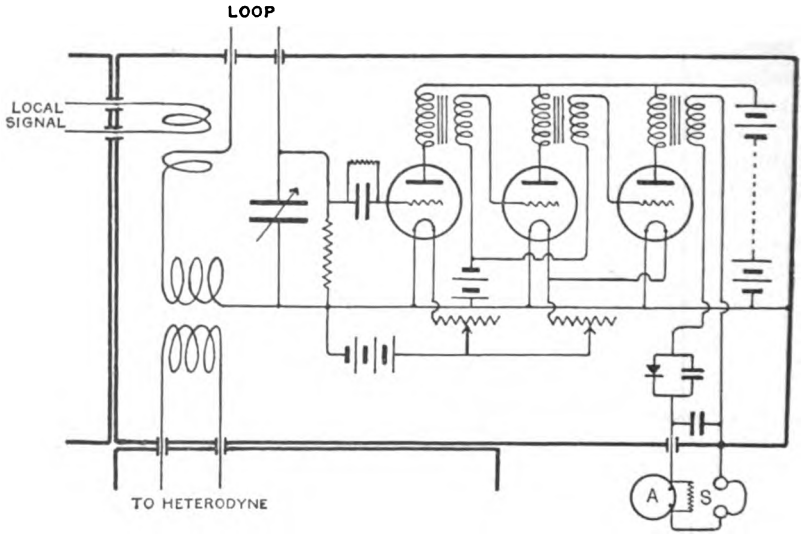


Fig. 3. Amplifier for 360 metres.

maximum intensity and the galvanometer shunt (Fig. 3, S) and the filament current in the amplifier are adjusted until a convenient reading is obtained on the galvanometer A. The aerial circuit is then accurately tuned and the frequency of the separate heterodyne is varied until a maximum reading is obtained on the galvanometer A. This maximum is very flat owing to the good characteristics of the transformers used. If the tuning of the aerial circuit is accurate the same reading is obtained with the heterodyne whistle on either side of the silent point. The loop is then turned round through 180° and the readings repeated. Any discrepancy in the two readings is probably a result of "antenna" effect and the mean is taken as the true reading. The heterodyne is left unaltered and the loop is turned through 90° to its position of minimum

Digitized by Google

intensity. The local signal is switched on and the variable resistance R set to an appropriate value. The frequency of the local signal is then varied till it gives a maximum deflection on the galvanometer A. This occurs when it is of exactly the same frequency as the London signals, which is easily perceived by means of the slow beats which are produced between the heterodyne whistle of the local signal and that due to the slight effect of London which always remains even at the minimum position. The strength of the local signal is then varied by varying the filament resistance until the reading obtained on the galvanometer A is the same as was observed for the London signals. The reading of the galvanometer B and the value of the resistance R then serve to determine the relative strength of the signal. The loop is then turned through 180° and the readings repeated. The mean is taken as the true value. Finally, a check reading is again taken on the London signals to verify that the amplifier has remained constant.

As has been mentioned above, the signal strength minimum obtainable on the loop was not perfect. This is probably due to "antenna effect." The value of the signal strength determined is the difference between the strengths at the position of maximum and minimum intensity. Provided, however, that the antenna effect remains constant the percentage of the signal which is cut out by the 90° rotation is always the same, and so the observed signal strength is still proportional to the London signal, which is all that is required. The antenna effect is of first order smallness and so a variation in it of first order will only produce a second order effect.

It was found, when using this method, that the modulation had no appreciable effect on the galvanometer readings.

We are informed by the B.B.C. that the transmission from 2 LO is constant to an accuracy of one per cent.

7. *Results.*

The results of the measurements on a wave length of 1600 m. are shown in the curve of Fig. 4. Here the product of signal strength and distance from the transmitter is plotted against the distance. The signal strengths used are, of course, only the relative strengths as obtained by the method described above. The continuous curve is drawn through the observed points and gives the experimental form of the function $f(\lambda, d, \rho, k)$. Mr D. Burnett of Clare College has very kindly calculated for us the corresponding values of the function f which are given by Sommerfeld's theory.

The crosses show points calculated for the value

$$\rho = 1{\cdot}8 \times 10^{13} \text{ E.M.U.}$$

Digitized by Google

At these wave lengths the dielectric constant is negligible in the formula and it is therefore not necessary to assume any definite value for it. It is seen that, for distances greater than 15 miles, the crosses fit on the curve well, and so $1{\cdot}8 \times 10^{13}$ E.M.U. may be taken as the value of ρ deduced from these results. The dotted curves have been drawn from calculations using the values $\rho = 1{\cdot}0 \times 10^{13}$ E.M.U. and $\rho = 2{\cdot}5 \times 10^{13}$ E.M.U. respectively, and they show the kind of accuracy with which ρ can be determined from the curve. It is seen that for distances less than about 13 miles the product of signal strength and distance decreases as we approach the transmitter. This is entirely contrary to Sommerfeld's theory, which predicts the dotted curves shown. The authors are

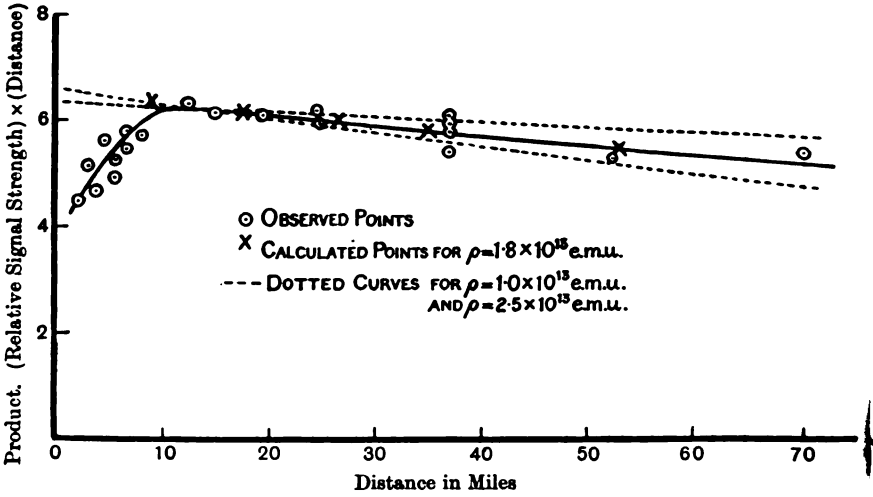


Fig. 4. Attenuation of 1600 m. waves.

unaware of any explanation of this effect, and are making further experiments on the attenuation at these very short distances. It is hoped to publish the results shortly.

Fig. 5 shows, in a similar way, the results obtained for the 360 m. wave. The observed points show a marked maximum near the 35 mile point. The contour of the ground is also shown in the figure and it seems obvious that the maximum is explained by the high ground near this point. In finding the ground attenuation we will neglect this maximum and use only the remaining points. The points used are indicated by circles, and the chain line of Fig. 5 has been drawn through them.

For this wave length the effect of diffractive bending becomes appreciable. The curve has accordingly been corrected for diffrac-

Digitized by Google

tive bending by employing the results given by Macdonald*. The corrected result is shown in the full-line curve of Fig. 5, which therefore represents the attenuation to be expected for transmission over a *flat* earth. The crosses in Fig. 5 are points calculated by Mr Burnett for $\rho = 0{\cdot}6 \times 10^{13}$ E.M.U. It is seen that they fit the curve fairly well. The dotted curves are drawn for $\rho = 0{\cdot}8 \times 10^{13}$ E.M.U. and $\rho = 0{\cdot}4 \times 10^{13}$ E.M.U. respectively. They serve to show the accuracy within which it is possible to estimate ρ. From this curve we are therefore led to the conclusion that $\rho = 0{\cdot}6 \times 10^{13}$ E.M.U. We also see that the accuracy with which ρ can be fixed is considerably greater than in the case of the 1600 m. signals.

The experiments give the two values $1{\cdot}8 \times 10^{13}$ and $0{\cdot}6 \times 10^{13}$ E.M.U. for ρ. These values agree fairly well when we remember the

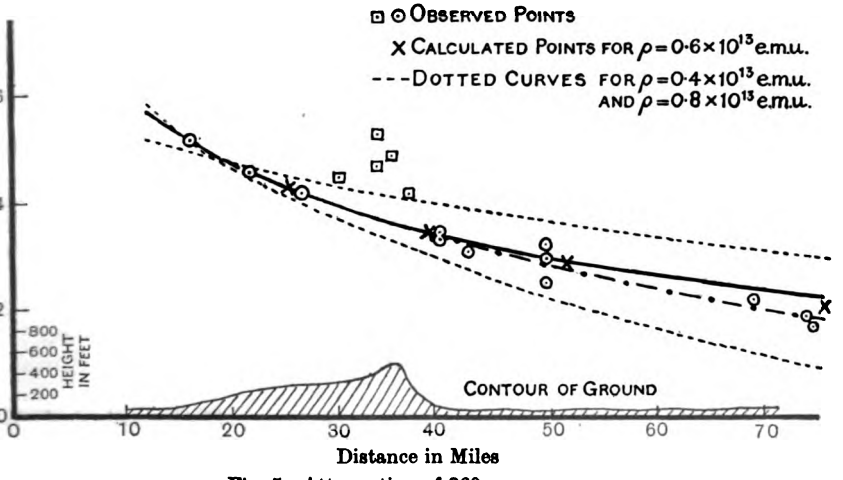


Fig. 5. Attenuation of 360 m. waves.

large change in wave length which has been made. They also lie within the limiting values $\rho = \cdot 19 \times 10^{13}$ E.M.U. and $\rho = 6{\cdot}7 \times 10^{13}$ E.M.U. found by Smith-Rose and Barfield† for various places in the south of England using an entirely different method. The average value given by Smith-Rose and Barfield is $\rho = 0{\cdot}5 \times 10^{13}$ E.M.U. which agrees with the value found for the 360 m. waves. The agreement of the values found for the two wave lengths with each other, and with the values found by other methods, serves to verify the theoretical expression of Sommerfeld.

The difference in the two values obtained for ρ may be due to the difference in the ground over which the measurements were made, or may be due to the change in the frequency of the waves.

* Macdonald, *Proc. Roy. Soc.* A, vol. CVIII (1925).
† *Exp. Wireless*, p. 745 (Sept. 1925).

Digitized by Google

Variation of attenuation with wave length. If we express the observed results approximately in the Austin-Cohen form

$$E = \frac{kh_r h_s i}{d} e^{-\alpha d/\sqrt{\lambda}}$$

we obtain for the coefficient α the values $\alpha = \cdot 003$ for 1600 m. and $\alpha = \cdot 006$ for 360 m., d and λ being measured in kilometres. It should be specially emphasised that these values for α are deduced from the parts of the curve corresponding to distances greater than 15 miles. If the curves depart appreciably from the exponential form for distances less than this, then the empirical formula will not give the true relation between the field strength at a very near point and at a distant point. For distances between 15 miles and 70 miles it is seen that a formula of this Austin-Cohen form does not give the correct change of attenuation with wave length.

We may also apply the results to the Fuller formula

$$E = \frac{kh_s h_r}{\lambda d} e^{-\alpha d/\lambda^{1\cdot 4}}.$$

This gives $\alpha = 0{\cdot}0028$ for the 360 m. signals and $\alpha = \cdot 0046$ for the 1600 m. signals. The value of α given by Fuller is $\cdot 0045$, which agrees with the present value for a wave length of 1600 m. The variation of α with wave length is here in a direction opposite to that obtained using Austin's formula, and so it would seem that if the change of attenuation with wave length is to be expressed in the form $e^{-\alpha d/\lambda^n}$, then n must lie between $\frac{1}{2}$ and 1·4. Such a formula, however, must be, at the best, only a rough approximation over a limited range.

Reflection coefficient of the Heaviside layer. If the value of α found above for the transmission from London to Cambridge is used in the formula* for calculating the reflection coefficient of the Heaviside layer, we arrive at a value of 30 per cent. for the reflection coefficient. Here again it is assumed that the attenuation law follows the exponential course right up to the transmitter.

8. *Accuracy of the Measurements.*

It is not claimed that these measurements are of any high accuracy. They are merely intended to give an idea of the order of magnitude involved. For high accuracy, in this kind of work, it is necessary to take large numbers of readings in order to eliminate local effects, and time and facilities were not available for this. The kind of accuracy to be expected has been indicated in paragraph 6.

* Appleton and Barnett, *Proc. Roy. Soc.* A, vol. CIX.

Digitized by Google

9. *Summary.*

1. It is pointed out that the signal strength of a wireless wave at a distant point depends on:

(*a*) the electrical constants of the ground,

(*b*) the curvature of the earth,

(*c*) the existence of an "atmospheric" ray coming downwards from the Heaviside layer.

2. We can eliminate the effect of (*b*) and (*c*), and so obtain direct evidence about the electrical constants of the ground, by making measurements near the transmitter.

3. Measurements on wave lengths of 300 m. and upwards give information about the resistivity of the ground, and on shorter wave lengths (15 m.) give the dielectric constant of the ground.

4. Attenuation measurements have been made over short distances for wave lengths of 1600 m. and 360 m.

5. The results are compared with those calculated from Sommerfeld's theory and show close correlation for distances beyond 10 wave lengths, but show deviations from the theory for shorter distances.

They give as values for the resistivity of the ground

$$\rho = 1{\cdot}8 \times 10^{13} \text{ E.M.U.}$$

for the Daventry signals, and

$$\rho = 0{\cdot}6 \times 10^{13} \text{ E.M.U.}$$

for the London signals.

In conclusion we wish to express our gratitude to Professor E. V. Appleton of King's College, London, for suggesting the problem and for his continual help and advice during the work. We wish also to thank Professor Sir Ernest Rutherford for affording facilities for carrying out the work, and we are especially indebted to Mr D. Burnett of Clare College for help on the mathematical side of the problem.

Note added 17 *June* 1926. Since this paper was written further measurements have been made within 20 miles of the London (2 LO) transmitter which have shown that calculations based on a single value of "ρ" will not give as close an agreement with the experimental results as can be obtained by assuming that ρ changes with the nature of the ground. For example the results up to 20 miles agree with the value $\rho = 2 \times 10^{13}$ E.M.U., whereas the results for greater distances, as shown above, indicate the value $\rho = 0{\cdot}6 \times 10^{13}$ E.M.U. It is of interest to note that such a change in ρ would be anticipated from the known geological nature of the ground, which consists of London Clays up to about 20 miles from the transmitter and of chalk at greater distances.

Digitized by Google

Some Relations between the Optical Spectra of Different Atoms of the same Electronic Structure. II. *Aluminium-like and Copper-like Atoms.* By Mr D. R. HARTREE, Fellow of St John's College.

[*Received* 22 May, *read* 26 July, 1926.]

§ 1. *Introduction.*

If it is assumed that the orbits of the series electron of an atom can be treated to a first approximation as orbits, specified by two quantum numbers n_k, in a central field of force, it is possible to derive theoretically approximate relations between corresponding terms of different atoms of the same electronic structure without knowing the details of the nature of the field of force of the atom. Some such relations, chiefly for lithium-like and sodium-like atoms, have been given by the writer in a previous paper*, where it is also shown that the actual terms of spectra of atoms of these types agree fairly well (in fact, rather better than might have been expected) with the theoretical approximate formulae†.

For these two types of atoms the relations are particularly simple, and also at the time of writing the previous paper they were the only types for which enough spectra were known to give a test of any theoretical formulae.

It is possible, however, to derive formulae which appear to be more general, and from the recent experimental work of Fowler, Millikan and Bowen, Carroll and others‡ there are now some data available on which to test these formulae. The establishment of these formulae and the comparison of them with experiment is the main subject of this paper.

§ 2. *Summary of Relevant Results of Previous Paper.*

The main results of the writer's previous paper will first be summarised for convenience of reference§. One or two of these results will be extended slightly for the purposes of the present paper.

* D. R. Hartree, *Proc. Roy. Soc.* vol. CVI, p. 552 (1924).

† For sodium-like atoms better data are now available for P V and S VI from Millikan's recent work (*Phys. Rev.* vol. XXV, p. 295 (1925)), and they confirm the agreement.

‡ A. Fowler, *Phil. Trans. Roy. Soc.* vol. CCXXV, p. 1 (1925) (Spectra Si II and III); R. A. Millikan and I. S. Bowen, *Phys. Rev.* vol. XXV, pp. 591, 600 (1925) (Spectra P III and IV, S IV and V); J. A. Carroll, *Phil. Trans. Roy. Soc.* vol. CCXXV, p. 357 (1926) (Spectra Ge III, Ge IV, In III, Hg II, Tl III, Pb IV); G. v. Salis, *Ann. der Phys.* vol. LXXVI, p. 145 (1924) (Spectra Zn II and Cd II); M. O. Saltmarsh, *Proc. Roy. Soc.* vol. CVIII, p. 332 (1925) (Spectra P III and IV).

§ Formulae quoted from the previous paper will be numbered with a Roman I prefixed to their number in that paper.

Digitized by Google

Electric charges are expressed as multiples of e, lengths as multiples of a, the radius of the 1_1 orbit of hydrogen, and electric potentials as multiples of the potential at distance a from a charge e. The atomic number is written N, the core charge C, and Z is written for the effective nuclear charge at any point, defined as the charge which, placed at the nucleus, would give the same electric intensity at that point as the electric intensity in the actual atomic field.

Structure of Atomic Field. It is convenient to write the potential v at any point radius ρ as the sum of the potential C/ρ due to a point charge equal to the core charge C, and an *added potential* $A(\rho)$ due to the finite size of the core. Let ρ_n be the average time mean radius of the core orbits in the group of principal quantum number n, and s_n the number of electrons in it; the added potential can be expressed, approximately at any rate, as the sum of contributions from the different groups of core orbits, in the form

$$A(\rho) = \sum_n \rho_n^{-1} \phi_n (\rho/\rho_n); \qquad \text{.....(I, 2·7)}$$

it is not necessary to know the exact behaviour of the functions ϕ, but their behaviour for small and for large values of the argument will be required. For x large, $\phi_n(x)$ tends to zero fairly rapidly [probably $\phi_n(x)$ is always $O(x^{-2})$, and is usually $O(x^{-4})$, for x large]; for x small

$$\phi_n(x) \sim s_n/x. \qquad \text{.....(I, 2·6)}$$

$\phi_n(x)$ is always less than its asymptotic value, and for the purposes of this paper an approximation to the difference of ϕ from its asymptotic value will be required; it will be assumed that

$$\phi_n(x) = (s_n/x) - s_n + O(x). \qquad \text{.....(2·1)}$$

For a group of s_n circular orbits radius ρ_n which have been averaged out into a spherically symmetrical shell of charge, this is correct without the $O(x)$ term; and as only an approximate value of the constant term is required, this probably gives a good enough estimate.

The maximum value of the principal quantum number n for core orbits will be written M.

Dimensions of Core Orbits. The average time mean radius ρ_n of a group of core orbits varies with the atomic number N, probably approximately in such a way that its reciprocal is linear in N. We will write

$$\rho_n (N - \gamma_n) = n^2 K_n \qquad \text{.....(2·2)}$$

[cf. formulae (I, 2·8, 6·41, 6·51)]; if the core orbits were all circular and the field were central, each K_n would be 1.

It may be noted here that the *dimensions* of the core orbits depend only on the *intensity* of the field in which the electron

Digitized by Google

moves, and so are not greatly affected by the addition of core orbits in groups of higher principal quantum number, which remain for the most part outside them. But the *energies* of the orbits depend on the *potential* there and so may be affected considerably by the presence of core orbits outside them.

Thus so long as the number of electrons in the groups of principal quantum number n and less remain the same, the mean radius ρ_n of the group of principal quantum number n can be taken, as a first approximation, to be independent of the degree of ionisation; and also the *dimensions* of the core orbits of given quantum number are much more nearly smooth functions of the atomic number than is the *energy*, and the curves of $1/\rho_n$ against N should not show breaks corresponding to stages in the building up of groups of higher principal quantum number such as the curves of $\sqrt{(\nu/R)}$ show*.

Quantum Defect and Term Excess. For the n_k term of an atom core charge C, the *quantum defect* q is defined by

$$\nu/R = C^2/(n-q)^2, \qquad \text{.....(I, 6·1)}$$

and for penetrating orbits the relations between corresponding terms of different atoms can often be best expressed in terms of q.

The variation of quantum defect within a sequence (k = const.) is most conveniently specified by Q defined by

$$Q = dq/d\,(\nu/R) \qquad \text{.....(2·3)}$$

[cf. (I, 6·32)]. If the sequence of terms fits the Rydberg-Ritz formula, Q is constant within the sequence.

For non-penetrating orbits, which lie so far outside the core that second order terms in the difference from the corresponding orbits of a hydrogen-like atom of the same core charge is negligible, it is more convenient to introduce the term excess α defined by

$$\nu/R = C^2/n^2 + \alpha\,(n, k); \qquad \text{.....(I, 3·1)}$$

to the first order the relation between q and α is

$$q = n^3\alpha/2C^2. \qquad \text{.....(2·4)}$$

Orbits nearly Hydrogen-like. If the added potential at a distance from the core can be expressed by a simple monomial formula

$$A\,(\rho) = \beta\rho^{-m}, \qquad \text{.....(I, 3·4)}$$

then for nearly hydrogen-like terms:

(1) for different terms of the same atom with the same value of k

$$\alpha\,(n, k).n^{m+1}/P_{m-2}\,(n/k) = \text{const.} \qquad \text{.....(2·5)}$$

[cf. I (3·5)], P_r being the Legendre polynomial of order r,

* N. Bohr and D. Coster, *Zeit. f. Phys.* vol. XII, p. 361 (1923).

Digitized by Google

(2) for different circular orbits of the same atom

$$\alpha\,(k, k).k^{2m} = \text{const.}, \qquad \ldots\ldots(\text{I}, 3{\cdot}8)$$

(3) for corresponding terms of different atoms

$$\alpha/\beta C^m = \text{const.} \qquad \ldots\ldots(\text{I}, 3{\cdot}9)$$

The index m in the added potential can be determined by comparison of (2·5) with experiment, and for sodium-like atoms and for C's the value $m = 4$ is obtained, which is the value to be expected if the external added field is due to the polarisation of the core by the series electron*.

Relation (I, 3·9) holds generally for any kind of atom; when atoms of the same electronic structure are considered, the variation of the coefficient β with atomic number is given by

$$\beta \propto {\rho_M}^{m-1}, \qquad \ldots\ldots(\text{I}, 3{\cdot}92)$$

ρ_M being the average time mean radius of the core orbits in the outermost group. From (I, 3·9) and (2·2) this gives, for corresponding terms of such atoms,

$$\alpha \propto C^m {\rho_M}^{m-1} \qquad \ldots\ldots(\text{I}, 3{\cdot}93)$$

$$\propto C^m/(N - \gamma_M)^{m-1}, \qquad \ldots\ldots(\text{I}, 3{\cdot}95)$$

or

$$(C^m/\alpha)^{1/(m-1)} \propto N - \gamma_M, \qquad \ldots\ldots(\text{I}, 3{\cdot}96)$$

so that $(C^m/\alpha)^{1/(m-1)}$ should be linear in the atomic number (or core charge).

This formula, with $m = 4$, agrees fairly closely with the relation between the observed d terms of different sodium-like atoms, and also, more closely, for the p terms of lithium-like atoms.

Penetrating Orbits. The quantum defect for the orbit of wave number ν in a central field of potential v is given by

$$2\pi q = \oint [2v - \nu/R - k^2/\rho^2]^{\frac{1}{2}}\, d\rho - \oint [2C/\rho - \nu/R - k^2/\rho^2]^{\frac{1}{2}}\, d\rho. \qquad \ldots\ldots(\text{I}, 6{\cdot}3)$$

Introducing (I, 2·7) for the added potential, and (2·2) for the average time mean radius of a group of core orbits, it is possible to obtain an asymptotic expression for the way in which q varies with atomic number N (or core charge C) for large values of the core charge.

* The agreement is best if integer values of k are adopted. In the writer's previous paper this agreement was advanced as evidence for integral values of k, but it seems probable now that the appropriate values of k to take are the half-integral values $\frac{1}{2}$, $\frac{3}{2}$, $\frac{5}{2}$ ... for s, p, d ... terms. The consequent divergence of the experimental values from the theoretical relation (2·5) may be due to the presence of further terms in the formula for the added potential, or to a variation of the coefficient β in (I, 3·4) from one orbit to another. This coefficient β is proportional to the polarisability of the core, and it was shown in (I, § 4) that the observations indicate a considerable variation of β with the k of the orbit of the series electron, whether integer or half-integer values of k are adopted.

Digitized by Google

For corresponding s terms of the spectra of lithium-like atoms, for which the core consists of the two K orbits ($n = 1$, $s_n = 2$) only, it was shown (I, § 1·6) that q tends asymptotically to be proportional to ρ_1, the radius of these orbits, so that $1/q$ tends asymptotically to be linear in the atomic number (see I, 6·44, 6·45, 6·46); and similar results also hold for the s and p terms of sodium-like atoms. Comparison with experiment shows that the asymptotic relations can be extended to small values of the core charge without great error (see I, § 7).

It was also shown that for the s orbits of sodium-like atoms, which penetrate first into the region of the L core orbits ($n = 2$) and also into the region of the K orbits ($n = 1$), the quantum defect can be expressed as the sum of contributions from these two groups of core orbits, and further that the contribution to the quantum defect from the K orbits is approximately equal to the quantum defect for the corresponding orbit of the lithium-like atom of the *same element* (see I, 6·63). The work on penetrating orbits which will be considered in this paper consists mainly in a generalisation of this result, and development of the consequences of the more general result.

Relations between the quantity $Q = dq/d\,(\nu/R)$, which measures the variation of quantum defect within a sequence, and core charge for corresponding sequences of different atoms of the same electronic structure can also be obtained. For s orbits of lithium-like atoms Q tends asymptotically to be proportional to ρ_1^3 (I, 6·71) so that Q/q^3 tends to a constant value (I, 6·73), and $(1/Q)^{\frac{1}{3}}$ tends to be linear in the core charge (I, 6·72).

Reference may be made here to a paper by Fues*, which is also concerned with the relations between spectra of different atoms of the same electronic structure, with special reference to penetrating orbits of lithium-, sodium-, copper-, and silver-like atoms. Fues writes the wave number of the n_k term in the form

$$\nu/R = (N - \sigma)^2/n^2, \qquad \text{.....(2·6)}$$

and for corresponding terms of different atoms of the same electronic structure finds a formula for σ of the form

$$\sigma = \sigma_0 + \frac{a}{N - b}, \qquad \text{.....(2·7)}$$

σ_0 being the limiting value of σ as $N \to \infty$, and a and b being constants for any one set of such terms.

This formula is first stated as an empirical one, and later in Fues' paper a theoretical justification of it for larger values of N is given. The present writer has shown independently† that if σ is

* E. Fues, *Ann. der Phys.* vol. LXXVI, p. 299 (1924).

† D. R. Hartree, *Proc. Camb. Phil. Soc.* vol. XXII, p. 464 (1924).

Digitized by Google

defined by (2·6), then for corresponding terms of different atoms of the same atomic structure, σ tends asymptotically to a constant value if the quantum defect q tends asymptotically to

$$\Sigma_n w_n/(N-\gamma_n')$$

as should theoretically be the case for penetrating orbits [see § 5 of this paper for this general formula], and suggested a formula equivalent to (2·7).

It may perhaps be mentioned that except for one or two numerical results specified in § 3, the work here considered is independent of the question whether integer or half-integer values for the quantum number k should be adopted.

§ 3. *Non-penetrating Orbits.*

For non-penetrating orbits no theoretical results for the relations between corresponding terms of different spectra have been obtained beyond those given in the previous paper (I, § 3), the most important of which are quoted here, but it may be pointed out that the relation

$$\alpha \propto C^m/(N-\gamma_M)^{m-1} \qquad \text{.....(I, 3·94)}$$

becomes in terms of quantum defect

$$q \propto C^{m-2}/(C+\kappa_M)^{m-1},$$

on using (2·4) and putting

$$\kappa_n = N - C - \gamma_n, \qquad \text{.....(3·1)}$$

so that κ_M is the number of core electrons less the screening number γ_M for the dimensions (not energy) of the outermost group of core orbits (κ_M is probably less than s_M, the number of electrons in this group). As C increases, it is clear that for large C, q decreases as $1/C$, but for small C the numerator may for a time increase more rapidly than the denominator, so that at first q increases with C, but later decreases. The point where the sign of the variation of q with C changes depends somewhat on the index m in the formula for the added potential, but chiefly on κ_M, occurring sooner the smaller κ_M. Actually for the d terms of sodium-like atoms q increases from Na I to S VI, but for p terms of lithium-like atoms, for which κ_M cannot be greater than 2, it only increases from Li I to Be II, and then decreases.

Fues (*loc. cit.*) states that for non-penetrating orbits q increases with C, without qualification.

For lithium-like and sodium-like atoms, the orbits either penetrate or remain so far outside the core that first order formulae are applicable; for atoms of many other types this is not always the case, often the innermost set of non-penetrating orbits (i.e. those with the smallest value of k) approach rather near the core,

Digitized by Google

and first order calculations cannot be applied to them, and usually terms corresponding to orbits lying further out are not at present known accurately enough to be of any use as data on which to test the first order formulae.

For Al I and Si II it is probable that term values are known fairly accurately*; some f terms are known for both these spectra and the first order formula for variation of term excess within the sequence (I, 3·5) can be tested on these terms. As shown in Table I, the observed values agree very fairly well with the calculated values of the term excess ratio for an inverse fourth power added potential, so it seems that for aluminium-like atoms also the external field is probably due to polarisation of the core of the series electron†.

Table I. *Term Excess Ratio for Aluminium-like Atoms.*

	$\frac{a(n,k)}{a(4,k)}$ obs., f terms		$\frac{a(n,4)}{a(4,4)}$ calc.		$\frac{a(n,3\frac{1}{2})}{a(4,3\frac{1}{2})}$ calc.
n_4	Al I	Si II	$m=3$	$m=4$	$m=4$
4_4	1·000	1·000	1·000	1·000	1·000
5_4	0·596	0·610	0·512	0·604	0·574
6_4	0·386	0·375	0·296	0·379	0·353

Polarisability deduced from f terms.

ϵ/a^3	Al^+	62	Si^{++}	31
	Al^{+++}	0·78	Si^{++++}	0·54

ϵ = Polarisability.
= Doublet strength per unit applied field.
a = Radius of 1_1 orbit of H atom.
ϵ/a^3 is a pure number.

From the values of the term excess, the value of the polarisability of the core can be calculated [cf. I (4·1)]‡; the large magnitude of the polarisability of the ions Al^+ and Si^{++} is interesting, and also its rapid decrease from Al^+ to Si^{++}. From a comparison of the polarisabilities calculated from the spectra Al I and Al III, it is clear that for Al I almost all the polarisation arises from the two large 3_1 orbits which lie (for the most part) well outside the neon-like structure which forms the core of the atom giving the

* A. Fowler, *Report on Series in Line Spectra* (Al I), *Phil. Trans.* (*loc. cit.*) (Si II).

† As for sodium-like atoms, the agreement with the values calculated for an inverse fourth power added potential is best if integer values of k are assumed. Cf. footnote, p. 307.

‡ The value $k=4$ has been used in determining the values of the polarisability tabulated; if the value $k=3\frac{1}{2}$ were used, all the values tabulated would be reduced by a factor of 2·18.

Digitized by Google

Al III spectrum. It would be expected that these orbits, being large and rather loosely bound, would be easily polarised, but even so the extent to which they actually seem to be polarised seems rather large. The rapid decrease of polarisability from Al I to Si II would also be expected, as the screening of the nuclear charge by the inner 10 electrons of the neon-like structure being probably almost complete in the outer part of the 3_1 orbits, these orbits would be expected to shrink rapidly as the atomic number was increased and much more rapidly than the inner group of 10, so that the ratio of the polarisabilities of Si^{++} and Al^{+} would be expected to be smaller than that of Si^{++++} and Al^{+++}, as in fact it is. Unfortunately no *f* terms of other aluminium-like atoms are yet known.

The *d* terms of aluminium-like atoms are so far from the terms of hydrogen-like atoms that first order formulae cannot be applied to them and at present there seems to be no simple way of treating theoretically non-penetrating orbits giving such terms. The variation of quantum defect and of doublet separation in the *d* sequence of Al I is somewhat abnormal, and so is the variation of quantum defect of the first *d* term with atomic number, for aluminium-like atoms.

§ 4. *Penetrating Orbits—General Expression for Quantum Defect.*

For penetrating orbits the chief new results are generalisations of the result that for *s* orbits of sodium-like atoms the quantum defect can be expressed as the sum of contributions from the different groups of core orbits into which the series electron penetrates, and that the contribution from the *K* group is approximately equal to the quantum defect of the corresponding orbit of the lithium-like atom of the same element.

If s_n is the number of core electrons in the group of principal quantum number n in an atom of atomic number N, ionised to such an extent that the core charge is C, and M is the maximum value of n for which there are core orbits, then

$$C = N - \sum_{n=1}^{M} s_n. \qquad \text{.....(4·1)}$$

The added potential can be represented approximately by

$$A(\rho) = \sum_{n=1}^{M} \rho_n^{-1} \phi_n(\rho/\rho_n), \qquad \text{.....(I, 2·7)}$$

where for small ρ/ρ_n it follows from (2·1) that

$$\rho_n^{-1} \phi(\rho/\rho_n) = s_n(1/\rho - 1/\rho_n) \qquad \text{.....(4·2)}$$

approximately, and for large ρ/ρ_n, $\phi_n(\rho/\rho_n)$ tends to zero fairly rapidly.

Digitized by Google

Let us write $$C_n = N - \sum_{u=1}^{n} s_u, \qquad \text{.....(4·3)}$$

so that C_n is the core charge of the atom of the given element, atomic number N, in such a state of ionisation that the core consists solely of the groups of core orbits of principal quantum number n or less, each of these groups being the same as it is in the atom in the state of ionisation for which the core charge is C. In particular

$$C = C_M = C_{M-1} - s_M. \qquad \text{.....(4·4)}$$

Also let us write

$$2\pi\xi_n(N, x) = \oint [2C_n/\rho + 2\rho_n^{-1}\phi_n(\rho/\rho_n) - x - k^2/\rho^2]^{\frac{1}{2}}\, d\rho - \oint [2C_n/\rho - x - k^2/\rho^2]^{\frac{1}{2}}\, d\rho, \qquad \text{.....(4·5)}$$

C_n being given by (4·3). C_n and ρ_n are functions of the atomic number N alone so long as the numbers of electrons in the groups of principal quantum number n and less do not alter; the ξ's are functions of k as well as of N and x, but orbits of different sequences will be considered separately, and the argument k will be suppressed. The function $\xi_n(N, x)$ is the quantum defect for the orbit of azimuthal quantum number k and wave number Rx, of an imaginary atom with core charge C_n already defined, but whose core orbits comprise *only* those of principal quantum number n, and are the same in number and dimensions of those of the actual atom.

Introducing (I, 2·7) into formula (I, 6·3) for quantum defect, we have

$$2\pi q = \oint [2C/\rho + 2\sum_{n=1}^{M} \rho_n\phi_n(\rho/\rho_n) - \nu/R - k^2/\rho^2]^{\frac{1}{2}}\, d\rho - \oint [2C/\rho - \nu/R - k^2/\rho^2]^{\frac{1}{2}}\, d\rho,$$

and using (4·4) this can be written in the form

$$\begin{aligned} 2\pi q = &\oint [2(C_{M-1} - s_M)/\rho + 2\sum_{n=1}^{M-1} \rho_n^{-1}\phi(\rho/\rho_n) \\ &+ 2\rho_M^{-1}\phi_M(\rho/\rho_M) - \nu/R - k^2/\rho^2]^{\frac{1}{2}}\, d\rho \\ &- \oint [2(C_{M-1} - s_M)/\rho + 2\rho_M^{-1}\phi_M(\rho/\rho_M) - \nu/R - k^2/\rho^2]^{\frac{1}{2}}\, d\rho \\ &+ \oint [2C_M/\rho + 2\rho_M^{-1}\phi_M(\rho/\rho_M) - \nu/R - k^2/\rho^2]^{\frac{1}{2}}\, d\rho \\ &- \oint [2C_M/\rho - \nu/R - k^2/\rho^2]^{\frac{1}{2}}\, d\rho. \qquad \text{.....(4·6)} \end{aligned}$$

Digitized by Google

The second pair of integrals can be put equal to $2\pi\xi_M(N, \nu/R)$ immediately. Also ρ_M is usually several times larger than ρ_{M-1} (especially for the lighter atoms), and $\phi_n(\rho/\rho_n)$ tends to zero fairly rapidly as ρ increases, for $\rho/\rho_n > 1$, so the term in ρ_n $(n < M)$ in the first integral is only appreciable for such values of ρ that the term in ρ_M has almost reached its asymptotic value $s_M(1/\rho - 1/\rho_M)$ given by (4·1). As the term in ρ_n $(n < M)$ is the only one in which the two integrals differ, we may substitute in them this asymptotic value for the term in ρ_M. Then (4·6) becomes

$$2\pi[q - \xi_M(N, \nu/R)]$$
$$= \oint [2C_{M-1}/\rho + 2\sum_{n-1}^{M-1} \rho_n^{-1}\phi_n(\rho/\rho_n) - (\nu/R + s_M/\rho_M) - k^2/\rho^2]^{\frac{1}{2}}\, d\rho$$
$$- \oint [2C_{M-1}/\rho - (\nu/R + s_M/\rho_M) - k^2/\rho^2]^{\frac{1}{2}}\, d\rho. \qquad \ldots\ldots(4\cdot7)$$

By subtracting

$$\oint [2C_{M-1}/\rho + 2\rho_{M-1}^{-1}\phi_{M-1}(\rho/\rho_{M-1}) - (\nu/R + s_M/\rho_M) - k^2/\rho^2]^{\frac{1}{2}}\, d\rho$$

from the first integral and adding it to the second, and writing $C_{M-1} = C_{M-2} - s_{M-1}$, we can put the right-hand side of (4·7) into the form of (4·6), and by repeating the argument as often as necessary we finally find

$$q = \sum_{n-1}^{M} \xi_n(N, \nu/R + \sum_{u-n+1}^{M} s_u/\rho_u). \qquad \ldots\ldots(4\cdot8)$$

Thus q is expressed as the sum of contributions from the different groups of core orbits, and the contribution of each group depends on the presence of groups outside it only through the term $\Sigma s_u/\rho_u$ in the second argument of (4·8), the effect of which will in most cases be small.

This expression is not exact, as both the formulae (I, 2·7) and (4·2) used in deducing it are only approximate; but the errors are probably not large.

This formula for the quantum defect as the sum of contributions from different groups of core orbits contains no specific reference to the fact that the field of the atom is assumed to be central, and it is possible that the formula is more general, just as the formula (I, 6·34) for $Q = dq/d(\nu/R)$ in terms of the radial periods of the actual and hydrogen-like orbits, deduced there on the assumption of a central field, contains no specific reference to that assumption and actually is more general.

§ 5. *Asymptotic Values of the Functions* ξ_n.

It is possible to find an asymptotic value for each $\xi_n(N, x)$ as $N \to \infty$, just as an asymptotic value was found in I, § 6 (formula

Digitized by Google

6·44) for q_0 for the s orbits of lithium-like atoms, which is in fact $\xi_1(N, 0)$ for $k = 1$.

An asymptotic value for $\xi_n(N, x)$ can be found either if x is regarded as constant as N varies, or if x varies in a stated way with N; in applications, x is the value of νR for an atom core charge C_n, and if for different atoms of the same core structure *corresponding* terms are taken, as is usually the case, x/C_n^2 remains finite and non-zero as N and C_n tends to infinity, so that it is for such a variation of x with N that an asymptotic value of $\xi_n(N, x)$ is required. Let us write

$$x = C_n^2 y,$$

and first consider y constant.

Substituting this, and putting $\rho/\rho_n = r$ [cf. I, 6·43], (4·5) becomes

$$2\pi\xi_n(N, C_n^2 y) = \oint [2C_n\rho_n/r + 2\rho_n\phi_n(r) - C_n^2\rho_n^2 y - k^2/r^2]^{\frac{1}{2}}\, dr$$

$$- \oint [2C_n\rho_n/r - C_n^2\rho_n^2 y - k^2/r^2]^{\frac{1}{2}}\, dr. \quad \ldots\ldots(5\cdot1)$$

Now so long as the numbers of core electrons in groups of principal quantum numbers n and less remain the same, $N - C_n$ is constant; so from (2·2) and (3·1)

$$(C_n + \kappa_n)\rho_n = n^2 K_n, \quad \ldots\ldots(5\cdot2)$$

where K_n and κ_n are nearly independent of C_n, and κ_n is indeed probably less than s_n, the number of core electrons with principal quantum number n. So as N and $C_n \to \infty$,

$$C_n\rho_n \to n^2 K_n. \quad \ldots\ldots(5\cdot3)$$

Hence if y is constant, the term $2\rho_n\phi_n(r)$ in (5·1) tends to zero as $C_n \to \infty$, while the other terms tend to finite limits, and it is the only term in which the two integrals differ; so for large C_n

$$\xi_n(N, C_n^2 y) \sim g_n(y)\rho_n, \quad \ldots\ldots(5\cdot4)$$

$$2\pi g_n(y) \doteq \oint \phi_n(r)\,[2n^2K_n/r - n^4K_n^2 y - k^2/r^2]^{-\frac{1}{2}}\, dr, \quad \ldots(5\cdot41)$$

$g_n(y)$ is independent of N or C_n, so that ξ_n tends asymptotically to become proportional to ρ_n. In this formula it must be remembered that n is the principal quantum number of a group of *core* orbits, and k is the azimuthal quantum number of an orbit of the *series* electron.

The same result holds if, instead of x/C_n^2 being constant and equal to y independent of C_n,

$$x/C_n^2 \to y \text{ as } C_n \to \infty. \quad \ldots\ldots(5\cdot42)$$

Digitized by Google

Substituting for ρ_n from

$$\rho_n (N - \gamma_n) = n^2 K_n, \qquad \text{.....(2·2)}$$

and writing

$$w_n (y) = n^2 K_n g_n (y), \qquad \text{.....(5·5)}$$

(5·4) becomes

$$\xi_n (N, C_n{}^2 y) \sim w_n (y)/(N - \gamma_n). \qquad \text{.....(5·6)}$$

But in deriving the asymptotic relation (5·4), quantities of the order γ_n/N have already been neglected compared to unity, so the denominator of this last result (5·6) might equally well be written N or C_n, and the γ_n is not really significant. For the extrapolation of the asymptotic relations to observed spectra, however, this is not good enough, since for these spectra the difference between N and C_n is never negligible; but without detailed analysis it can be seen that a formula of the type (5·6) may be expected to hold approximately, with some quantity γ_n' other than γ_n substituted for γ_n.

Consider ξ_n expressed as two terms of a power series in ρ_n, and a remainder, by substituting for $C_n \rho_n$ from (5·2) in (5·1) and expanding the two integrals in powers of ρ_n; the first term in ξ_n is obviously $g_n (y)\, \rho_n$; let us drop the arguments of ξ_n and g_n for shortness, and write

$$\xi_n = g_n \rho_n + (b_n/n^4 K_n{}^2)\, \rho_n{}^2 + O (\rho_n{}^3) \qquad \text{.....(5·7)}$$

(the coefficient of the quadratic term has been put in this form for convenience); b_n depends on y, and is not zero when y is zero*.

Multiplying (5·7) by $(N - \gamma_n)$ and using (2·2), (5·5) gives

$$\begin{aligned} \xi_n (N - \gamma_n) &= w_n + b_n/(N - \gamma_n) + O [(N - \gamma_n)^{-2}] \\ &= w_n + b_n \xi_n / w_n + O (\xi_n{}^2), \end{aligned}$$

or

$$\xi_n (N, C_n{}^2 y) = w_n (y)/[N - \gamma_n' (y)] + O (\xi_n{}^2), \qquad (5·8)$$

$$\gamma_n' (y) = \gamma_n (y) + b_n (y)/w_n (y).$$

Thus to the first order $1/\xi$ should be linear in N (or C_n), but the constant γ_n' in (5·8) is different from the γ_n of (2·2), by an amount which depends on y, and is not zero when y is zero.

Neglecting the variation of the ξ's and W's with y, (4·8) can be written

$$q \sim \sum_{n=1}^{M} w_n/(N - \gamma_n'), \qquad \text{.....(5·9)}$$

approximately, by using (5·8).

The contribution $\rho_n{}^{-1} \phi_n (\rho/\rho_n)$ to the added potential from the groups of core orbits with principal quantum number n depends on s_n, the number of electrons in the group; for given ρ and ρ_n, it is, in fact, approximately proportional to s_n. So if two types of atom which have different numbers s_n of electrons in this group

* If $x/C_n{}^2$ is not constant but tends to y as $N \to \infty$, b_n contains a term depending on $[x/C_n{}^2 - y]/\rho_n$; i.e. on how $x/C_n{}^2$ tends to y as $N \to \infty$.

Digitized by Google

are compared, g_n is approximately proportional to s_n by (5·41); also it is probable that K_n is nearly but not quite the same for the two types, so w_n should also be approximately proportional to s_n.

§ 6. *Penetrating Orbits—Relations between Values of Quantum Defect in Different Spectra of a Single Atom, and in Spectra of Different Atoms.*

We can express (4·8) in a more convenient manner for comparison with values of q deduced from actual spectra.

Consider the spectrum of an atom of atomic number N and core charge C, for which the outermost core orbits have principal quantum number M. The quantum defect q for the term of the spectrum of this atom wave number ν, in the sequence of azimuthal quantum number k, is by (4·8)

$$q = \sum_{n=1}^{M} \xi_n (N, \nu/R + \sum_{u=n+1}^{M} s_u/\rho_u). \qquad \text{.....(6·1)}$$

Consider also the spectrum of the atom of the same element ionised to such an extent that the groups of core orbits with principal quantum number $M-1$ and less are the same as in the atom core charge C, but there are no core orbits of principal quantum number M; the core charge of this atom is $C_{M-1} = C + s_M$, and the quantum defect q' for the term of wave number ν', in the sequence of same azimuthal quantum number k, is

$$q' = \sum_{n=1}^{M-1} \xi_n (N, \nu'/R + \sum_{u=n+1}^{M-1} s_u/\rho_u). \qquad \text{.....(6·2)}$$

Since each ρ_n is only affected very slightly by the presence of orbits of greater principal quantum number, the functions $\xi_n (N, x)$ in (6·1) and (6·2) can be treated as identical.

If we define the value of ν' to be used in (6·2) by

$$\nu'/R = \nu/R + s_M/\rho_M, \qquad \text{.....(6·3)}$$

then (6·1) can be written

$$\begin{aligned} q &= \sum_{n=1}^{M-1} \xi_n (N, \nu/R + \sum_{u=n+1}^{M-1} s_u/\rho_u + s_M/\rho_M) + \xi_M (N, \nu/R) \\ &= q' + \xi_M (N, \nu/R), \end{aligned}$$

or

$$q - q' = \xi_M (N, \nu/R). \qquad \text{.....(6·4)}$$

If now for different elements we compare corresponding terms, principal quantum number $\mathbf{n}$, of the spectrum of the less ionised of the two atoms (quantum defect q), then

$$\nu/R = C^2/(\mathbf{n} - q)^2,$$

and [since $C = C_M$, (4·4)]

$$(\nu/R)/C_M^2 \to 1/\mathbf{n}^2 \text{ as } C_M \to \infty.$$

Digitized by Google

Condition (5·42) is therefore satisfied, so (5·4), (5·7), (5·8) are applicable to $\xi_M(N, \nu/R)$ for such terms, and (6·4) leads to

$$q - q' \sim g_M \rho_M \sim w_M/(N - \gamma_M'). \qquad \ldots\ldots(6{\cdot}5)$$

These results are not yet very convenient for comparison with observed spectra, as the calculation of ν' involves the estimation of ρ_M, which cannot be determined at all accurately, and also the evaluation of q' requires interpolation between values of the quantum defect for the observed terms of the spectrum of more highly ionised atom. These difficulties can be avoided as follows.

If ν'' is the wave number of the term of the second atom with quantum defect q', then

$$\nu''/R = (C + s_M)^2/(n - q')^2, \qquad \nu/R = C^2/(n - q)^2,$$

and ν'' is almost certain to be greater than ν. The value of ν' defined by (6·3), which is the wave number corresponding to the orbit whose quantum defect q' is required in the formulae (6·4)–(6·5), is also greater than ν. It is not easy to see in general whether ν'' is a good approximation to ν', but for different atoms of the same electronic structure the error made in substituting q'' for q' must vary regularly and probably slowly with C, so it probably will not affect the form of the relation (6·5), though it may affect the values of the constants.

Thus for (6·5) we may substitute

$$q - q'' \sim g_M \rho_M \sim w_M/(N - \gamma_M'), \qquad \ldots\ldots(6{\cdot}6)$$

the constants g_M, w_M and γ_M' being probably slightly different from those in (6·5).

In words, consider two atoms, which we will call the X-like and Y-like atoms, of a single element in different states of ionisation such that the cores of the two atoms differ only in this, that the group of highest principal quantum number in the core of the Y-like atom is lacking in the X-like atom. Take the difference of the quantum defects for corresponding orbits of the Y-like and X-like atom of the same element (both being penetrating orbits); then if for different elements of the same electronic structure corresponding terms are taken, this difference of the quantum defect tends asymptotically to become proportional to the mean radius of the orbits of maximum principal quantum number in the core of the Y-like atom, and the reciprocal of the difference of the quantum defects tends to become linear in the atomic number (or core charge).

This result may be made clearer by an example.

The core of an aluminium-like atom consists of the core of the sodium-like atom of the same element, with core orbits with $n = 1$ and 2, and two 3_1 orbits; in this case ρ_3 is the time mean radius

Digitized by Google

of these two orbits and X and Y of the above general statement are Na and Al.

If q and q'' are values of the quantum defect for corresponding terms of spectra of the aluminium-like and sodium-like atoms of the same element, then for corresponding terms of aluminium-like atoms of different elements

$$q - q'' = \xi_3 (N, \nu/R)$$
$$\sim g_3 \rho_3$$
$$\sim w_3/(N - \gamma_3'), \qquad \text{.....(6·7)}$$

g_3 and w_3 being independent of N. It is necessary that the orbits of both atoms should penetrate, in order for these results to hold.

The quantities ξ_3, g_3, w_3, γ_3' in these formulae for aluminium-like atoms refer to the group of two 3_1 orbits in the core, and are different from the corresponding quantities for the 8 M orbits of a potassium-like atom, or the 18 M orbits of a copper-like atom.

§ 7. *Variation of Quantum Defect within a Sequence.*

Results similar to those of § 4 and § 6 for the quantum defect q can be obtained in a similar way for the quantity $Q = dq/d\,(\nu/R)$, which measures the variation of quantum defect within a sequence.

A general result, corresponding to (4·8) for q, can be obtained by differentiating this formula directly with respect to ν/R, and analysis similar to that of §§ 5, 6 can be carried out. The final results, corresponding to (6·5), (6·6), can be expressed as follows. If Q and Q'' are the values of $dq/d\,(\nu/R)$ for atoms of the same element in the two different states of ionisation defined in § 6, then for different atoms of the same electronic structure

$$Q - Q'' \sim G_M \rho_M^3 \qquad \text{.....(7·1)}$$
$$\sim W_M/(N - \gamma_M'')^3, \qquad \text{.....(7·2)}$$

G_M, W_M and γ_M'' being constants for a given group of core orbits; γ_M'' is different from γ_M' occurring in the similar formula for $q - q''$.

From (7·2), $(Q - Q'')^{\frac{1}{3}}$ tends asymptotically to become linear in N, and from (6·5) and (7·1) $(Q - Q'')/(q - q'')^3$ tends asymptotically to become constant. These are the generalisations of the results already found for sodium-like atoms (I, 6·72, 6·73).

§ 8. *Comparison with Experiment—Aluminium-like Atoms.*

Of the spectra of aluminium-like atoms, Al I and Si II are known well enough for series limits to be determined fairly accurately, so that for these spectra reliable values of the terms and quantum defects can be obtained; for P III and S IV some data are available and values of the quantum defect can be obtained

Digitized by Google

which are less certain than those for Al I and Si II, but not too uncertain to make a comparison with the theoretical formula worthless*.

The theoretical results are as follows. For an aluminium-like atom, the number of electrons s_3 in the outermost group of core orbits ($n = 3$) is 2; let ρ_3 be the time mean radius of these orbits. Let q be the quantum defect for the n_k term of an aluminium-like atom; and for the sodium-like atom of the same element let q'' be the quantum defect for the corresponding term. Then from the results of § 6

$$q - q'' \sim g_3\rho_3$$
$$\sim w_3/(N - \gamma_3') \qquad \text{.....(6·7)}$$

for C large, where g_3 and w_3 are constants; so that as C increases, $(q - q'')/\rho_3$ should tend to a constant value and $1/(q - q'')$ should tend to become linear in C†. Also if Q is the value of $dq/d\,(\nu/R)$ for a sequence of an aluminium-like atom and Q'' is the value of $dq/d\,(\nu/R)$ for the corresponding sequence of the sodium-like atom of the same element, then as C increases $1/(Q - Q'')^{\frac{1}{3}}$ should tend to become linear in C, and $(Q - Q'')/(q - q'')^3$ should tend to a constant value.

The 4_1, 5_1, 3_2 and 4_2 are the only terms corresponding to penetrating orbits for which values of q can at present be obtained reliable enough to give a test of these results. The values of q and derived quantities for testing these formulae are given in Table II. The uncertainty in term values for P III and S IV should be kept in mind in comparing these results with the theoretical formulae.

It is not possible to obtain any reliable estimate of ρ_3, the time mean radius of the two 3_1 core orbits; the quantity ρ_m tabulated is the maximum radius of the normal (3_1) orbit of the sodium-like atom of the same element, calculated from the corresponding term (formula I, 7·1; cf. I, Table VII), and to a first approximation it is probable that the ratio ρ_m/ρ_3 can be taken as constant.

In the case of lithium-like and sodium-like atoms it was found that the theoretical asymptotic relations for large values of the core charge still held fairly closely even for atoms of small core

* The values of the quantum defect tabulated for P III and S IV have been found from the term values given by Millikan and Bowen (*loc. cit.*). For P III Miss Saltmarsh (*loc. cit.*) calculates term values from the observed lines by assuming that the quantum defects for the first two s terms are the same, and obtains $2s = 127004$; the present writer has attempted to obtain a more accurate value by estimating the difference of quantum defect between the first two s terms from the value of the corresponding difference in Al I and Si II, and obtained $2s = 125560$ with an estimated limit of error of 150. Millikan and Bowen's value is $2s = 125498$.

† If ν is the wave number of the term in question, it is theoretically more accurate to replace q'' in these formulae by q', the quantum defect for the orbit of the sodium-like atom, with the same k, which would give a term of wave number $\nu' = \nu + 2R/\rho_3$ (cf. formulae 6·3–6·6).

Digitized by Google

charge: and the same is the case for the aluminium-like atoms. For each n_1 orbit $1/(q-q'')$ increases by approximately the same amount from one spectrum to the next, and $(q-q'')\rho_m$ appears to be rapidly tending to a constant value*; also for the s terms $1/(Q-Q'')^{\frac{1}{2}}$ increases by about the same amount from one atom to the next, and $(Q-Q'')(q-q'')^3$ remains approximately constant.

Table II. *Comparison of Quantum Defects for Penetrating Orbits of Aluminium-like Atoms.*

	Spectrum	Al I	Si II	P III	S IV
4_1	q	1·8125	1·4275	1·1945	1·0379
	q''	0·9677	0·7873	0·6972	0·6271
	$q-q''$	0·9048	0·6402	0·4973	0·4108
	$1/(q-q'')$	1·105	1·562	2·011	2·434
	Diffs.	457	449	423	
	ρ_m	2·692	2·279	1·987	1·767
	$(q-q'')\rho_m$	0·3361	0·2809	0·2502	0·2324
4_1–5_1	Δq	0·0313	0·0281	0·0255	0·0235
	$\Delta(\nu, R)$	0·1125	0·2959	0·5304	0·8163
	$Q=\Delta q\,\Delta(\nu, R)$	0·2785	0·0950	0·0481	0·0288
	Q''	0·0153	0·0095	0·0063	0·0043
	$Q-Q''$	0·2632	0·0855	0·0418	0·0245
	$1/(Q-Q'')^{\frac{1}{2}}$	1·56	2·27	2·88	3·44
	Diffs.	71	61	56	
	$(Q-Q'')(q-q'')^3$	0·355	0·326	0·340	0·353
3_2	q	1·4906	1·1732	0·9830	0·8521
	q''	0·6279	0·5493	0·4887	0·4406
	$q-q''$	0·8627	0·6239	0·4943	0·4115
	$1/(q-q'')$	1·159	1·603	2·023	2·430
	Diffs.	444	420	407	
4_2	q	1·3233	1·0541	0·8856	0·7655
	q''	0·6059	0·5281	0·4688	0·4218
	$q-q''$	0·7174	0·5260	0·4168	0·3437
	$1/(q-q'')$	1·394	1·901	2·399	2·909
	Diffs.	507	498	510	
	$(q-q'')/\rho_m$	0·2664	0·2306	0·2097	0·1945
	$\dfrac{(q-q'')\,4_1}{(q-q'')\,4_2}$	1·261	1·217	1·193	1·195

* These results hold equally well on the whole but not appreciably better when the theoretically more accurate q' is substituted for the practically more convenient q''; as it happens the effect of the difference between q' and q'' is not large and, for a given term, it varies little from atom to atom.

Digitized by Google

It seems probable that if the maximum radius ρ_3 of the 3_1 orbits were known, $(q - q'')/\rho_3$ would be found to be more nearly constant than the tabulated quantity $(q - q'')/\rho_m$. ρ_m is the maximum radius of the one 3_1 orbit of the sodium-like atom of the element in question, and it seems probable that the addition of a second 3_1 orbit would increase the maximum and mean radius of the first 3_1 orbit, and moreover would increase them in a greater ratio for atoms of small core charge C than for atoms of large C, so that ρ_3/ρ_m would be expected to decrease with increasing C. Since $(q - q'')/\rho_m$ also decreases with increasing C, $(q - q'')/\rho_3$ will be more nearly constant than $(q - q'')/\rho_m$*.

For the 3_2 orbits the field at aphelion probably differs considerably from that of a point charge C, as the aphelion distance of the 3_2 orbits is probably not much greater than that of the 3_1 orbits already present in the core, and the analysis from which the theoretical relations were deduced cannot be applied to such orbits; however the formulae actually hold as closely for these orbits as for the others.

It is interesting to note the relative values of the contributions of the outer two 3_1 core orbits and those of the rest of the core to q and Q. In the case of the quantum defect q, $q - q''$ is the contribution from the two 3_1 orbits and q'' the contribution from the rest of the core. In all cases as the core charge rises, the contribution from the outer two orbits decreases compared to that from the rest of the core, as would be expected since the dimensions of the outer core orbits alter proportionally more than those of the inner orbits with increasing core charge, and the contribution to q is proportional to the dimensions of the group of orbits. For these atoms, the contributions to q are about equal for the atoms of core charge 1 or 2.

The contribution of the outer two orbits to Q is a much greater proportion of the whole than is their contribution to q, as would be expected since the contribution to Q varies as the cube of the dimensions of the orbits while the contribution to q varies as the first power of the dimensions.

For any one orbit of the series electron, $q - q''$, the contribution to the quantum defect from the two 3_1 core orbits, tends asymptotically to be proportional to the mean radius ρ_3 of these orbits, so the ratio of values of $q - q''$ for two orbits of the series electron of different k tends asymptotically to a constant value; moreover, the deviations from the exact proportionality between $q - q''$ and ρ_3 are probably of the same sign and same order of magnitude for different orbits of the series electron, so for atoms of small core charge the constancy of the ratio of $q - q''$ for two

* It is unlikely that the variation of ρ_3/ρ_m is large enough to counteract entirely the decrease of $(q - q'')/\rho_m$.

Digitized by Google

orbits of different k would be expected to hold more closely than the other relations established theoretically as asymptotically true for large C. The ratio of values of $q - q''$ for the 4_1 and 4_2 orbits is given in the last line of the table, and its proportional variation is in fact much less than that of $(q - q'')/\rho_m$ for either orbit.

§ 9. *Other Atoms of Aluminium Sub-group Types.*

The neutral atoms of B, Al, Ga, In and Tl, and atoms of other elements ionised to such an extent that they have the same number of electrons as one of these neutral atoms, all have cores consisting of 2 electrons in n_1 orbits in addition to complete groups of electrons with smaller values of n. Let q be the quantum defect of a term (corresponding to a penetrating orbit) of one of these atoms, and q'' the quantum defect for the corresponding term of the atom of the same element ionised two stages further, i.e. without the two outermost n_1 orbits; then $q - q''$ is the contribution of the two outermost orbits to q. Further, the mean radius ρ_n of these orbits can be taken as approximately proportional to ρ_m, the maximum radius of the one n_1 orbit which is the normal orbit of the series electron of the atom ionised two stages further, and this can be calculated from the corresponding term of the spectrum of this atom.

Though the mean radius ρ_m of the two outer n_1 orbits might be different for two atoms, their contribution to the added charge at radius ρ would be expected to be much the same function of ρ/ρ_m for all atoms of aluminium sub-group types; i.e. for this pair

Table III. *Comparison of Contributions to Quantum Defect from outermost two n_1 Orbits of Atoms of Different Aluminium Sub-group Types.*

Atom	Al	Ga	In	Tl	C	Si
Spectrum	III	III	III	III	IV	IV
Normal orbit n_k	3_1	4_1	5_1	6_1	2_1	3_1
Normal orbit $n - q$	2·075	1·997	2·092	1·995	1·937	2·204
Normal orbit ρ_m	2·69	2·49	2·76	2·48	1·74	2·30
Spectrum	I	I	I	I	II	II
n_k	4_1	5_1	6_1	7_1	3_1	4_1
q	1·813	2·844	3·782	4·805	0·659	1·427
q''	0·908	1·962	(2·848)	(3·960)	0·157	0·787
$(q - q'')$	0·905	0·882	0·934	0·845	0·502	0·640
$(q - q'')/\rho_m$	0·336	0·352	0·338	0·341	0·289	0·281

Values in brackets are estimated; for In III and Tl III only the first s term is known.

Digitized by Google

of orbits ϕ_n in formulae (I, 2·7), (4·5), (5·4), (5·41) is probably nearly the same function of its argument for all values of n. Hence it would be expected that for large values of the core charge $(q-q'')/\rho_n$, and so $(q-q'')/\rho_m$, would be nearly the same, not only for aluminium-like atoms, but for all atoms of aluminium sub-group types.

Further, the deviations for small C from the asymptotic relations for large C are likely to be much the same for different atoms of the same C, so that though the result of these deviations is that for aluminium-like atoms $(q-q'')/\rho_m$ is not constant for small values of C, it may be expected that there will be little difference between the values of this ratio for B I, Al I, Ga I, In I, Tl I, or between the values for C II, Si II, Ge II, Sn II, Pb II, and so on. Table III shows that this is indeed the case.

§ 10. *Copper-like Atoms.*

For copper-like atoms there is enough reliable data now available to provide a test of the theoretical results; values of the quantum defect and derived quantities required are tabulated in Table IV*.

Table IV. *Comparison of Values of Quantum Defect for Penetrating Orbits of Copper-like Atoms.*

Spectrum	Cu I	Zn II	Ga III	Ge IV
4_1 q	2·6729	2·2594	2·0036	1·8178
$1/q$	0·3741	0·4426	0·4991	0·5501
Diffs.	685	565	510	
q''	0·2743	0·2629	0·2524	0·2428
$q-q''$	2·3986	1·9965	1·7512	1·5750
$1/(q-q'')$	0·4169	0·5009	0·5710	0·6310
Diffs.	840	701	639	
4_2 q	2·1342	1·8565	1·6636	1·5162
$1/q$	0·4686	0·5386	0·6011	0·6595
Diffs.	701	625	584	
q''	0·1835	0·1758	0·1688	0·1623
$q-q''$	1·9507	1·6807	1·4948	1·3539
$1/(q-q'')$	0·5126	0·5950	0·6690	0·7386
Diffs.	824	740	696	
4_3 q	[1·0211]†	0·9752	0·9126	0·8579
$1/q$		1·0254	1·0958	1·1656
Diffs.		704	698	

* The values of q have been calculated from the term values given by A. Fowler for Cu I (*Report on Series of Line Spectra*) by G. v. Salis for Zn II (*loc. cit.*) and by J. A. Carroll for Ga III and Ge IV (*loc. cit.*).

† If the first d term corresponds to a penetrating orbit.

Digitized by Google

The core of a copper-like atom consists of electrons in orbits with $n = 1$, 2 and 3. The electrons with $n = 1$ and 2 (K and L shells) together form a structure which is the core of the sodium-like atom of the same element; the quantum defect q for any term corresponding to a penetrating orbit is the sum of contributions from all three groups of core orbits, the contribution from the inner two groups together is very nearly equal to q'', the quantum defect of the corresponding term of the sodium-like atom of the same element (see § 6). Theoretically $1/(q - q'')$ for such terms should tend to become linear in the core charge C as C increases.

The spectra of 6 sodium-like atoms are now known fairly accurately and the values of q fit the theoretical relations well, so the values of q'' required in considering copper-like atoms can be extrapolated with some confidence, although the range of extrapolation is rather large*.

The contribution of the K and L shells to q for both s and p terms of these atoms is very small, about a tenth of the whole (the proportion increases with increasing core charge, as would be expected theoretically; cf. discussion of this point in connection with Al-like atoms), and it varies in the same sort of way as the contribution of the M shell, so for a given term the variation of $1/q$ with C is very similar to that of $1/(q - q'')$. Both quantities are tending to become linear in C and the departure from linearity with C (measured by the ratio of the second difference to the first difference) is almost the same for each.

The series d terms of the four spectra are interesting. For each spectrum, the first d term may correspond either to a 3_3 orbit lying well outside the core, in which case the term would be expected to have a small positive q (though negative values of q are perhaps not impossible), or it may correspond to a 4_3 orbit which penetrates the core, in which case it would be expected that it would have a q differing from that for the 4_2 orbit by about the same amount as the difference between the values of q for penetrating p and d orbits of other spectra (about 0·85 for $C = 2$, more for $C = 1$ and less for $C = 3$, but no other certain data are yet available). The term values for Zn II, Ga III and Ge IV strongly suggest penetrating orbits, and also $1/q$ is nearly linear in C for them as would be expected for penetrating orbits. The Cu I term might also correspond to a 4_3 orbit, with a value of q shown in square brackets in the table, but this seems unlikely, as it is clear from a glance at the figures that this value of q does not form a smooth sequence with the others; it seems very much more likely that for Cu I the first d term corresponds to a 3_3 orbit lying entirely

* The third decimal place in the values of q'' given in the table is probably reliable; the method of extrapolation used is such that the figures are *smooth* to the four places given, though the fourth figure is probably not correct.

Digitized by Google

outside the core, with $q = 0{\cdot}0216$. From a consideration of doublet separations Carroll* also concludes that the d terms of Cu I correspond to orbits of a different type from those corresponding to the terms of other atoms of the same electronic structure.

§ 11. *Other Atom Types.*

Some data for silver-like and gold-like atoms are available, though term values are probably subject to greater uncertainty than those for copper-like atoms; values of $1/q$ for the terms available have been calculated by Carroll†, and they suggest that, as for other atoms, $1/q$ tends to become linear in C as C increases.

For silver-like atoms the contribution of the outer group of core orbits (N shell, $n = 4$) to the quantum defect of the outer group would be found by subtracting from it the quantum defect for the corresponding term of the copper-like atom of the same element; not enough spectra of copper-like atoms are known to enable the required extrapolation to be carried out at all accurately, but a rough extrapolation suggests that the total contribution from the K, L, M groups is about a fifth of the whole quantum defect for the 5_1 and 5_2 terms of Ag I.

For both silver-like and gold-like atoms there appears to be a difference in type between the orbits corresponding to the d terms of the arc spectrum and the spark spectra, similar to the difference already mentioned for the copper-like atoms.

The theoretical relations also hold for the spectra of magnesium-like atoms; here, however, the question is somewhat complicated by the existence of terms of different multiplicities; the application of the relations to these spectra, and to non-series terms such as p' terms of spectra of magnesium-like and aluminium-like atoms, will be treated in a further paper.

§ 12. *Summary.*

The relevant results from the writer's previous paper on the relation between spectra of atoms of different atomic structure are summarised.

For non-penetrating orbits no new theoretical results are obtained, and there are few known spectra (other than those of lithium-like or sodium-like atoms already treated in the first paper) on which to test the relations previously obtained. Values of the polarisability for the Al^+ and Si^{++} ions are calculated from terms of Al I and Si II respectively corresponding to non-penetrating orbits, and are shown to be very much greater than the values of the polarisability of the neon-like ions Al^{+++} and Si^{++++}.

The main new results are those for penetrating orbits. Assuming a central field of force, it is shown that the quantum defect q for

* *Loc. cit.* § 3. † *Loc. cit.* Tables VI and VII.

Digitized by Google

such an orbit can be expressed as the sum of contributions from the electrons in groups of core orbits of different principal quantum number n, and further that if for a given atom in different states of ionisation corresponding orbits of the series electron are compared, the contribution to q from a set of core orbits of given n is very nearly independent of the degree of ionisation so long as the number of electrons in core orbits in the group remains the same. It follows that, if q is the quantum defect for a term of the spectrum of an atom core charge C, the core of which contains s_M orbits of principal quantum number M and none of higher quantum number, and q'' is the quantum defect for the corresponding term in the spectrum of the atom of the same element with core charge $C + s_M$, which differs from the atom of core charge C only in lacking the core orbits of principal quantum number M, then $q - q''$ is approximately the contribution from the core orbits of principal quantum number M to the quantum defect for the term of the atom core charge C.

Further, it is shown that if corresponding terms of different atoms of the same electronic structure are compared, then for large values of C the contribution to q from any group of core orbits should tend asymptotically to be proportional to the average time mean radius of these orbits, and its reciprocal should tend asymptotically to be linear in C.

Somewhat similar relations are obtained for the quantity $Q = dq/d(\nu/R)$ which measures the variation of quantum defect within a sequence.

These theoretical results, and in particular the result that $1/(q - q'')$ should tend asymptotically to be proportional to C, are compared with the values of q deduced from the terms of such observed spectra of aluminium-like and copper-like atoms as are available, and it is found that though the theoretical relations are only established as asymptotically true for large C, there is a considerable measure of agreement with spectra for small values of C, which are the only ones which can be observed.

Digitized by Google

On the Quantum Theory of Vibration-Rotation Bands. By J. R. OPPENHEIMER. (Communicated by Mr R. H. FOWLER.)

[*Received* 24 May, *read* 26 July, 1926.]

1. The frequencies of the lines in the vibration-rotation spectrum of a diatomic molecule may be obtained unambiguously by the application of the classical quantum dynamics to a simple molecular model. The dominant term in the intensities of these lines has been obtained by Fowler* from the summation rules; this, and the next term, were computed by Kemble† from an application of the correspondence principle to a three-dimensional model. But the details of these calculations are not free from ambiguity. For, in the first case, the weight of the states, and, in the second, the method of averaging over transitions, is arbitrary. For the outer lines of the band this ambiguity is negligible; for the inner ones, and particularly for the missing central line, it is not.

Now the new mechanics ‡ gives a precise solution of this problem, and might accordingly be expected to decide definitely the intensity of the null line, and the value of the higher terms in the intensities of the outer lines. The purpose of this note is to obtain the solution, and to compare it with previous results, and with the results of experiment.

The problem has been treated by Brillouin§. Brillouin's chief conclusions are (1) that, contrary to the contention of Heisenberg||, the new mechanics does not necessarily give the same energies as the classical theory with the introduction of half-integral quantum numbers; and (2) that, if Heisenberg's result were correct, the new mechanics would account for the disappearance of the central line. He uses Cartesian coordinates $(q_1 q_2)$, and their corresponding momenta $(p_1 p_2)$. He imposes the constraint

$$q_1^2 + q_2^2 = a^2,$$

and the single quantum condition

$$[q_1 p_1] + [q_2 p_2] = 1 \tag{1}$$

for the rotational motion, where $[ab]$ is the Poisson Bracket of

* Fowler, *Phil. Mag.* vol. XLIX, p. 1272 (1925).
† Kemble, *Phys. Rev.* vol. XXV, p. 1 (1925).
‡ Dirac, *Proc. Roy. Soc.* A, vol. CX, p. 561 (1926); Born, Heisenberg, Jordan, *Zeit. f. Physik*, vol. XXXV, p. 557 (1926).
§ Brillouin, *C.R.* vol. CLXXXII, p. 374 (1926).
|| Heisenberg, *Zeit. f. Physik*, vol. XXX, p. 879 (1925); Born und Jordan, *Zeit. f. Physik*, vol. XXXIV, p. 858 (1925).

Digitized by Google

a and b. On the other hand, the Heisenberg-Dirac* quantum conditions reduce, in this case, to a single condition

$$\left[(q_1 p_2 - q_2 p_1), \quad \frac{q_1 + iq_2}{(q_1^2 + q_2^2)^{\frac{1}{2}}}\right] = i\,\frac{q_1 + iq_2}{(q_1^2 + q_2^2)^{\frac{1}{2}}}$$

which cannot be reduced to the form (1). Moreover, in the Hamiltonian used by Brillouin the coupling of the kinetic energies of rotation and vibration is neglected. Finally, the problem is treated in two dimensions, so that we should not expect† correct values for the intensities. Brillouin does not give his solution in detail. He does give the intensities for the case of vanishing coupling, and for this case gets the same intensities for all lines of a band. But it may readily be shown that the differential equation is regular, uniformly in the coupling parameter λ as this tends to zero. It follows that the argument cannot explain the absence of the central line for the small finite values of λ which occur. In fact we shall show that, in the two-dimensional case, the central line would be present.

2. *The dynamical problem.* We may take for model two point-atoms attracting each other radially. If the masses of the atoms be m_1 and m_2, then the classical Hamiltonian, after the elimination of the centre of gravity, becomes

$$H = \frac{1}{2\mu}(p_x^2 + p_y^2 + p_z^2) + U_1(x^2 + y^2 + z^2), \tag{2}$$

where

$$\frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2}; \quad x = x_1 - x_2; \tag{3}$$

$$(m_1 + m_2)\,p_x = m_2 p_{x_1} - m_1 p_{x_2}.$$

Since we shall not ultimately want $x_1 y_1 z_1$, $x_2 y_2 z_2$, and since the elimination of the centre of gravity (3) holds in the quantum theory, we may take (2) as Hamiltonian. There is as yet no adequate reason for not choosing some classically equivalent form. The only other result which we shall take over directly from the classical theory is the expression for the energy radiated per second. We shall use this in terms of Cartesian coordinates and accelerations‡; and we shall not be inconsistent if we take the Hamiltonian in terms of them.

In the classical solution it is convenient to use polar angle variables. These cannot be represented by matrices, and we shall

* Dirac, *Proc. Roy. Soc.* A, vol. cx, p. 561 (1926).
† Kemble, *Phys. Rev.* vol. xxv, p. 1 (1925).
‡ Heisenberg, *Zeit. f. Physik*, vol. xxxiii, p. 879 (1925); Born und Jordan, *Zeit. f. Physik*, vol. xxxiv, p. 858 (1925).

Digitized by Google

accordingly adopt the more general methods developed by Dirac*. We want a canonical transformation to uniformizing variables such that xyz shall be periodic in the angle variables. Classically we should obtain this by three successive transformations: introduction of polar coordinates, in which the Hamilton-Jacobi equation separates; elimination of the nodes; and uniformization of the vibration-rotation. Dirac† has shown that there is a canonical transformation equivalent to the first two:

$$(4)\quad r^2 = x^2+y^2+z^2;\quad p_r = \frac{x}{r}p_x + \frac{y}{r}p_y + \frac{z}{r}p_z - \frac{ih_*}{r},$$

$$\epsilon^{i\vartheta} = r^{-1}k^{\frac{1}{2}}(k_2^2 - p^2)^{-\frac{1}{4}}(z + i[kz])\,k^{\frac{1}{2}};$$

$$k^2 = \tfrac{1}{4}h_*^2 + \Sigma_x m_x^2 = \tfrac{1}{4}h_*^2 + \Sigma_x (yp_z - zp_y)^2,$$

$$\epsilon^{i\phi} = (k^2 - p_2^2)^{-\frac{1}{2}}(m_x + im_y);\quad p = xp_y - yp_x,$$

where $\quad p_2 = p - \tfrac{1}{2}h_*, \qquad k_2 = k - \tfrac{1}{2}h_*, \qquad h_* = \dfrac{1}{2\pi}h.$

[In general we write $\alpha_1 = \alpha + \frac{1}{2}h_*$; $\alpha_2 = \alpha - \frac{1}{2}h_*$, if α is an action variable.] We shall use only the following information about this transformation:

1. It is canonical.

2. The values of xyz obtained by the inversion of (4) are:

$$(5)\quad z = \tfrac{1}{2}r(\varpi_2\epsilon^{i\vartheta} + \varpi_1\epsilon^{-i\vartheta}),$$

$$\left.\begin{matrix} x \\ iy \end{matrix}\right\} = \frac{r}{4i}[\varpi_{22}\epsilon^{i\vartheta+i\phi} \pm \varpi_{21}\epsilon^{i\vartheta-i\phi} - \varpi_{12}\epsilon^{-i\vartheta+i\phi} \pm \varpi_{11}\epsilon^{-i\vartheta-i\phi}],$$

where $\quad \varpi_1^2 = 1 - \dfrac{p_1p_2}{k(k+h_*)}; \qquad \varpi_2^2 = 1 - \dfrac{p_1p_2}{k(k-h_*)};$

$$\varpi_{11}^2 = \frac{(k+p_1)(k+p_1+h_*)}{k(k+h_*)}.$$

$$(6)\quad \varpi_{12}^2 = \frac{(k-p_2)(k-p_2+h_*)}{k(k+h_*)}; \quad \varpi_{21}^2 = \frac{(k-p_1)(k-p_1-h_*)}{k(k-h_*)};$$

$$\varpi_{22}^2 = \frac{(k+p_2)(k+p_2-h_*)}{k(k-h_*)}.$$

3. The transformed Hamiltonian is

$$(7)\qquad H = \frac{1}{2\mu}\left(p_r^2 + \frac{k_1k_2}{r^2}\right) + U(r).$$

The third transformation requires certain hypotheses about U; in analogy with the classical solution‡ we shall suppose that the

* Dirac, *Proc. Roy. Soc.* A, vol. cx, p. 561 (1926); Born, Heisenberg, Jordan, *Zeit. f. Physik*, vol. xxxv, p. 557 (1926).

† Dirac, *Proc. Roy. Soc.* A, vol. cxi, p. 281 (1926).

‡ Born, *Atommechanik*, § 20.

Digitized by Google

molecule can rotate uniformly; that this rotation is stable; and that the amplitude of the oscillations about it which do in fact occur is small compared with the mean distance of separation of the atoms. Then we must have

$$\dot{p}_r = [p_r H] = 0; \quad \dot{r} = [rH] = 0; \quad k_1 k_2 = \mu r^3 [U'(r), p_r]$$

when we substitute the proper quantum numbers for k, and the corresponding c-number for r. We shall take over this c-formula, which may be justified physically, and write

$$(8) \qquad \frac{k_1 k_2}{\mu a^3(k)} = [U p_r]_{r=a(k)} = U'(a) \text{ (Def.)},$$

where k and a are q-numbers. Further, we set $r = a(k) + \rho$, and thus obtain from (1) and (8), the expansion

$$(7\,a) \qquad H = \frac{k_1 k_2}{2\mu a^2(k)} + \frac{1}{2\mu}(p_r^2 + \omega^2 \rho^2) + \lambda \rho^3 - \ldots$$

which converges by hypothesis. In (7 a)

$$\omega^2 = \mu U''(a) + \frac{3 k_1 k_2}{a^4(k)} = \omega_0^2 + \frac{3 k_1 k_2}{a^4(k)}; \quad \lambda = \tfrac{1}{6} U'''(a) - 4 \frac{k_1 k_2}{2\mu a^5}.$$

We may note further that

$$[\rho p_r] = 1, \qquad [\rho k] = 0,$$

but that

$$(9) \qquad [\rho \vartheta] = -[a \vartheta] = \frac{2k}{\omega^2 a^3(k)}.$$

The retention of first order terms in λ does not affect the frequencies, but changes the intensities. It follows, however, from the form of the transformation for r and ϑ, that the relative intensities of lines in the same band will be changed only by a factor of the order of $k^2 J^{\frac{1}{2}}/a^5 \omega^{\frac{5}{2}}$, which is small compared with the terms we shall retain, and too small for experimental detection. The new mechanics is, of course, applicable to the more general case; but the transformation, even in the simplest case, where $U'''(a) \sim 0$, is very cumbersome. We shall accordingly neglect λ and observe that our results will not give us the intensities in the overtone bands.

The transformation

$$(10) \quad r = \frac{1}{i(2\omega)^{\frac{1}{2}}}(J^{\frac{1}{2}} \epsilon^{iw} - \epsilon^{-iw} J^{\frac{1}{2}}) + a; \quad p_r = \{\tfrac{1}{2}\omega\}^{\frac{1}{2}}(J^{\frac{1}{2}} \epsilon^{iw} + \epsilon^{-iw} J^{\frac{1}{2}}),$$

$$k = k, \qquad \vartheta = \tau + \alpha;$$

$$\alpha = -\left(\frac{2}{\omega}\right)^{\frac{1}{2}} \cdot \frac{k}{a^3 \omega}(J^{\frac{1}{2}} \epsilon^{iw} + \epsilon^{-iw} J^{\frac{1}{2}}) - \frac{3k}{4a^4 \omega^2 i}$$

$$\times \{J^{\frac{1}{2}}(J-h)^{\frac{1}{2}} \epsilon^{2iw} - \epsilon^{-2iw} J^{\frac{1}{2}}(J-h)^{\frac{1}{2}}\},$$

$$p = p, \qquad \phi = \phi$$

Digitized by Google

defines the new canonical variables (J, ω), (k, τ), (p, ϕ). If we assume that these satisfy the quantum conditions, we get at once

$$[kp] = [kp_r] = [pp_r] = [\phi p_r] = [\phi \Im] = [\phi k] = [\phi \gamma] = 0;$$
$$[p\phi] = [k\Im] = -1.$$

Moreover

$$(11) \quad [rp_r] = -[p_r r] = \frac{1}{ih_*}\frac{1}{2i}(2J - 2\epsilon^{-iw}J\epsilon^{iw}) = 1,$$

$$(12) \quad i(\rho p_r + p_r\rho) = (J-h)^{\frac{1}{2}}J^{\frac{1}{2}}\epsilon^{2iw} - \epsilon^{-2iw}J^{\frac{1}{2}}(J-h)^{\frac{1}{2}}$$

and

$$(13) \quad p_r^2 + \omega^2\rho^2 = \tfrac{1}{2}\omega(2J + 2\epsilon^{-iw}J\epsilon^{iw}) = 2\omega J_1.$$

From the first of these (11) it follows that r and p_r are canonical. From the second (12) we get, neglecting powers of $\frac{k\sqrt{\mu}}{a^2\omega}$ above the second,

$$[\Im r] = -\frac{3k\mu\rho}{2a^4\omega^2} + \frac{3k\mu}{4a^4\omega^2}[p_r\rho + \rho p_r, \rho] - \frac{2k}{a^3\omega^2} + \frac{2}{\omega}\frac{k}{a^3\omega}[rp_r] = 0,$$

and

$$[\Im p_r] = +\frac{3k\mu p_r}{2a^4\omega^2} - \frac{3k\mu}{4a^4\omega^2}[p_r\rho + \rho p_r, p_r] = 0.$$

Assuming therefore that the new variables satisfy the quantum conditions, we have shown that the old variables satisfy all the quantum conditions. This establishes the transformation. The asymmetry of r and p_r, and thus of W, in J, is not peculiar to this problem. It has its analogue in the harmonic oscillator, for which

$$|q(n, n+1)|^2 \propto n+1.$$

3. *Frequencies, Quantum Numbers.* By (7 *a*) and (13)

$$(14) \quad W = W(Jk) = \frac{k_1k_2}{2\mu a^2(k)} + \frac{\omega J_1}{\mu}.$$

This differs from the classical value in having J_1 for J, and k_1k_2 for k^2.

Before we can find the frequencies, we must decide how to assign the quantum numbers, n, m, s, corresponding to the action variables J, k, p. The hypothesis of a normal state for the vibratory motion means that n must be integral. The argument by which m and s are determined is strictly analogous to that given by Dirac* and Born, Heisenberg, and Jordan† for the Zeeman effect of an atom with one series electron. We should expect s to be so chosen that

$$\left|\frac{p_1p_2}{k(k-h_*)}\right| \leqslant 1; \qquad \left|\frac{p_1p_2}{k(k+h_*)}\right| \leqslant 1,$$

since, otherwise by (5) z might not be real for real s and m. This means

* Dirac, *Proc. Roy. Soc.* A, vol. CXI, p. 281 (1926).

† Dirac, *Proc. Roy. Soc.* A, vol. CX, p. 561 (1926); Born, Heisenberg, Jordan, *Zeit. f. Physik*, vol. XXXV, p. 557 (1926).

Digitized by Google

that s may vary from $-m+\frac{1}{2}$ to $+m-\frac{1}{2}$; it is readily seen from the values we get for the intensities (18), that all transitions from s values within this range to those outside are forbidden. Finally, we may choose m to give a normal state for which a transition with decreasing m is impossible. In this case the only physical justification for such an assumption is the observed absence of the central line*. We shall show that the central line does not vanish if the only condition on m is Brillouin's condition of symmetry for positive and negative rotations.

With this assignment of quantum numbers (clearly s is irrelevant) we get for the absorption frequencies

(14 a)

$$\nu(n, m \to n+1, m \pm 1) = \frac{h}{8\pi^2\mu a^2(0)}\left(\mp 2m + 1 - \frac{h^2(m \pm \frac{1}{2})^3}{\pi^2 a^4(0)\,\omega_0{}^2}\cdots\right) + \nu_1(n+\tfrac{3}{2}) - \nu_2(n+\tfrac{1}{2}),$$

$$\nu_1 = \frac{\omega(m \pm 1)}{2\pi\mu} \sim \frac{\omega_0}{2\pi\mu}\left(1 + \frac{3\mu h^2(m \pm 1 + \frac{1}{2})(m \pm 1 - \frac{1}{2})}{2a^4(0)\,\omega_0{}^2}\right);$$

$$\nu_2 \sim \frac{\omega_0}{2\pi\mu}\left(1 + \frac{3\mu m_1 m_2 h^2}{2a^4(0)\,\omega_0{}^2}\right).$$

The difference between these and the classical frequencies is too small for experimental detection, more particularly since $a(0)$ and ω_0 can only be found from the band itself. The energy differences would, however, give an appreciable term in the specific heats.

4. *Intensities.* We may now compute the *a priori* intensities of the band lines directly. For this we need x, y, z in terms of the final variables. In each case we have several alternative forms. Thus

(15)

$$z = \frac{r}{2}(\varpi_2\epsilon^{i\beta} + \epsilon^{-i\beta}\varpi_2) = \tfrac{1}{2}(\epsilon^{i\beta}\varpi_1 + \epsilon^{-i\beta}\varpi_2)\,r = \frac{r}{2}(\varpi_2\epsilon^{i\beta} + \varpi_1\epsilon^{-i\beta}) \ldots.$$

But since all the transformations are canonical, these forms ought to give the same intensities; and it may be readily verified that they do. Now to the second order in $\frac{h}{a^2\omega_0}$, we get, using the expansion for the exponential

(16)
$$\epsilon^{ir+i\alpha} = \sum_0^\infty{}_n \frac{1}{n!}(i\alpha + i\tau)^n \sim \sum_0^\infty{}_n \frac{(i\tau)^n}{n!} + \sum_0^\infty{}_n \sum_0^{n-1} \frac{(i\tau)^r (i\alpha)(i\tau)^{n-1-r}}{n!}$$

$$= \sum_0^\infty{}_n \frac{n(i\alpha)(i\tau)^{n-1}}{n!} + \sum_0^\infty{}_n \frac{(i\tau)^n}{n!} - \sum_0^\infty{}_n \frac{n(n-1)\,ih_*(i\tau)^{n-2}\alpha}{2kn!}$$

$$= \sum_0^\infty{}_n \left(1 + \frac{i\alpha.k_2}{k_2}\right)\frac{(i\tau)^n}{n} = \left(1 + \frac{i\alpha k_2}{k}\right)\epsilon^{ir} = (1 + i\alpha_2)\,\epsilon^{ir},$$

$$\epsilon^{-ir-i\alpha} = \left(1 - \frac{i\alpha.k_1}{k}\right)\epsilon^{-ir} = (1 - i\alpha_1)\,\epsilon^{-ir}.$$

* On Schrödinger's Theory there is no normal state for a rigid rotator in two dimensions, and the transitions for m, $(\frac{1}{2} \to -\frac{1}{2})$, $(-\frac{1}{2} \to \frac{1}{2})$ occur.

Digitized by Google

Then from (5), (10) and (16),

$$(17)\quad z = \tfrac{1}{2}\left[a(k) + \frac{1}{i(2\omega)^{\frac{1}{2}}}(J^{\frac{1}{2}}\epsilon^{iw} - \epsilon^{-iw}J^{\frac{1}{2}})\right]\times [\varpi_2(1+i\alpha_2)\epsilon^{ir} + \varpi_1(1-i\alpha_1)\epsilon^{-ir}],$$

$$\left.\begin{matrix} x \\ iy \end{matrix}\right\} = \frac{1}{4i}\left[a(k) + \frac{1}{i(2\omega)^{\frac{1}{2}}}(J^{\frac{1}{2}}\epsilon^{iw} - \epsilon^{-iw}J^{\frac{1}{2}})\right][\varpi_{22}(1+i\alpha_2)\epsilon^{ir+i\phi} \pm \varpi_{21}(1+i\alpha_2)\epsilon^{ir-i\phi} - \varpi_{12}(1-i\alpha_1)\epsilon^{-ir+i\phi} \mp \varpi_{11}(1-i\alpha_1)\epsilon^{-ir-i\phi}].$$

Now expand $\omega^{-\frac{1}{2}}$ to second order in $\dfrac{k}{a^2\omega}$, take α_1 and α_2 in front of the exponentials in w, and collect the coefficients of ϵ^{iw} and ϵ^{-iw}. This gives for the Zeeman components

(18)

$$|z(-1,-1,0)| = \frac{1}{2}\left(\frac{n}{2\omega_0}h_*\right)^{\frac{1}{2}}\varpi_2[1-2\gamma m_2\{1+\tfrac{3}{8}\gamma(m_2-n)\}],$$

$$|z(-1,+1,0)| = \frac{1}{2}\left(\frac{n}{2\omega_0}h_*\right)^{\frac{1}{2}}\varpi_1[1+2\gamma m_1\{1-\tfrac{3}{8}\gamma(m_1+n)\}],$$

$$|z(+1,-1,0)| = \frac{1}{2}\left(\frac{n+1}{2\omega_0}h_*\right)^{\frac{1}{2}}\varpi_2[1+2\gamma m_2\{1-\tfrac{3}{8}\gamma(m_2+n+1)\}],$$

$$|z(+1,+1,0)| = \frac{1}{2}\left(\frac{n+1}{2\omega_0}h_*\right)^{\frac{1}{2}}\varpi_1[1-2\gamma m_1\{1+\tfrac{3}{8}\gamma(m_1-n-1)\}];$$

$$|x(-1,-1,-1)| = \frac{1}{4}\left(\frac{nh_*}{2\omega_0}\right)^{\frac{1}{2}}\varpi_{22}[1-2\gamma m_2\{1+\tfrac{3}{8}\gamma(m_2-n)\}],$$

$$|x(-1,-1,+1)| = \frac{1}{4}\left(\frac{nh_*}{2\omega_0}\right)^{\frac{1}{2}}\varpi_{21}[1-2\gamma m_2\{1+\tfrac{3}{8}\gamma(m_2-n)\}],$$

$$|x(-1,+1,-1)| = \frac{1}{4}\left(\frac{nh_*}{2\omega_0}\right)^{\frac{1}{2}}\varpi_{12}[1+2\gamma m_1\{1-\tfrac{3}{8}\gamma(m_1+n)\}],$$

$$|x(-1,+1,+1)| = \frac{1}{4}\left(\frac{nh_*}{2\omega_0}\right)^{\frac{1}{2}}\varpi_{11}[1+2\gamma m_1\{1-\tfrac{3}{8}\gamma(m_1+n)\}],$$

$$|x(+1,-1,-1)| = \frac{1}{4}\left(\frac{n+1}{2\omega_0}h_*\right)^{\frac{1}{2}}\varpi_{22}[1+2\gamma m_2\{1-\tfrac{3}{8}\gamma(m_2+n+1)\}],$$

$$|x(+1,-1,+1)| = \frac{1}{4}\left(\frac{n+1}{2\omega_0}h_*\right)^{\frac{1}{2}}\varpi_{21}[1+2\gamma m_2\{1-\tfrac{3}{8}\gamma(m_2+n+1)\}],$$

$$|x(+1,+1,-1)| = \frac{1}{4}\left(\frac{n+1}{2\omega_0}h_*\right)^{\frac{1}{2}}\varpi_{12}[1-2\gamma m_1\{1+\tfrac{3}{8}\gamma(m_1-n-1)\}],$$

$$|x(+1,+1,+1)| = \frac{1}{4}\left(\frac{n+1}{2\omega_0}h_*\right)^{\frac{1}{2}}\varpi_{11}[1-2\gamma m_1\{1+\tfrac{3}{8}\gamma(m_1-n-1)\}];$$

$$|y(ijk)| = |x(ijk)|; \quad |x(ijk)| = |x[(n,m,s)\to(n+i,m+j,s+k)]|,$$

where $$\gamma = \frac{h}{2\pi\omega_0 a^2(0)} = \frac{h}{4\pi^2\mu a_0^2.\nu_0} = \frac{\nu_r}{\nu_0}.$$

Digitized by Google

Actually the system is degenerate. Summing the "intensities"

$$I_0(0, m \to 1, m+1) = \text{const.}\,[1 - 4\gamma m_1 (1 - \tfrac{3}{8}\gamma m_1 - \tfrac{3}{8}\gamma)]\,\bar{S},$$

$$\bar{S} = \sum_{-m+\frac{1}{2}}^{+m-\frac{1}{2}} s\,\varpi_1^2 + \tfrac{1}{2}\sum_{-m+\frac{1}{2}}^{+m-\frac{3}{2}} (s)\,\varpi_{11}^2 + \tfrac{1}{2}\sum_{-m+\frac{3}{2}}^{+m-\frac{1}{2}} (s)\,\varpi_{12}^2$$

$$= \tfrac{1}{2}\sum_{-m+\frac{1}{2}}^{+m-\frac{1}{2}} (\varpi_{11}^2 + 2\varpi_1^2 + \varpi_{12}^2)$$

$$= \tfrac{1}{2}.m.\left(4 + \frac{2}{m}\right) = 4\,(m + \tfrac{1}{2}) = 4m_1.$$

Similarly

$$I_0(0, m+1 \to 1, m) = 4.\text{const.}\,[1 + 4\gamma m_1 (1 + \tfrac{5}{8}\gamma m_1 - \tfrac{3}{8}\gamma)].m_1.$$

The intensity of the central line vanishes. But it would not have vanished if we had made m integral, nor if we had treated the problem in two dimensions.

For the intensities of the $\bar{n}$th lines* (in energy, and not in quanta)

$$I_+(\bar{n}) \propto \bar{n}\,(1 - 4\gamma\bar{n}\,(1 - \tfrac{5}{8}\gamma\bar{n} - \tfrac{3}{8}\gamma))\,\epsilon^{-\frac{1}{kT}W(0,\bar{n}-\frac{1}{2})}.\nu(0, \bar{n} - \tfrac{1}{2} \to 1, \bar{n} + \tfrac{1}{2}),$$

$$I_-(\bar{n}) \propto \bar{n}\,(1 + 4\gamma\bar{n}\,(1 + \tfrac{5}{8}\gamma\bar{n} - \tfrac{3}{8}\gamma))\,\epsilon^{-\frac{1}{kT}W(0,\bar{n}+\frac{1}{2})}.\nu(0, \bar{n} + \tfrac{1}{2} \to 1, \bar{n} - \tfrac{1}{2}),$$

or, at least for not too large m,

$$I_+(\bar{n}) \propto \bar{n}\,\epsilon^{-\frac{1}{kT}W(0,\bar{n})}\,(1 - 3\gamma\bar{n})\left(1 + \frac{\bar{n}h^2}{2\mu a_0^2 k\tau}\right),$$

$$I_-(\bar{n}) \propto \bar{n}\,\epsilon^{-\frac{1}{kT}W(0,\bar{n})}\,(1 + 3\gamma\bar{n})\left(1 - \frac{\bar{n}h^2}{2\mu a_0^2 k\tau}\right).$$

This in part† resolves the paradox noted by Fowler. For the two branches are nearly equal in intensity in spite of the assymmetry of the temperature factor. The main part of the assymmetric factor which compensates this had been previously obtained by Kemble from the Coriolis forces.

This means that the summation rules do not apply exactly, because the frequency differences of the lines of the band are not negligible. Neglecting the factors $(1 \pm 4\gamma\bar{n} \ldots)$ and $\epsilon^{-\frac{1}{kT}W}$, they hold exactly. For the weights of the states are proportional to m.

I should like to thank Mr R. H. Fowler, F.R.S., and Mr P. A. M. Dirac for their criticism and advice.

* Kemble, *loc. cit.* and Fowler, *Phil. Mag.* vol. L, p. 1079 (1925).

† For the fundamental hydrogen chloride band studied by Kemble and Fowler the values of 4γ and $h^2/2\mu a_0^2 k\tau$ differ by only about 10 per cent.; the positive branch remains slightly the stronger.

Digitized by Google

5. *Summary.* The dynamical problem of the "diatomic molecule" is solved on the new mechanics. The terms of the rotational energy are $A\,(m^2 - \frac{1}{4})$, where $m = \pm\frac{1}{2}, \frac{3}{2}, \frac{5}{2} \ldots$; the weights of the corresponding states are $2m$; the frequencies differ a little from the classical ones. Finally the intensities are slightly different from those computed by Kemble; the main term agrees with that of Fowler, but the positive branch is only slightly stronger than the negative. The central line vanishes. The intensities are valid only for the fundamental band.

6. *Addendum.* This problem has been treated by Schrödinger* by his method of characteristics. Schrödinger does not find the intensities; but he obtains energy terms which give frequencies in agreement with (14) and (14 *a*). More recently the problem has been treated by Mensing† by the method of matrices in a paper which came to hand just as the writing of this paper was completed. Mensing neglects γ but retains λ. The main terms in the intensities of the fundamental band are, of course, the Zeeman factors of (18). By an application of Born and Jordan's solution for the anharmonic oscillator‡ Mensing also obtains expressions for the intensities in the overtones. It is, however, readily seen from (17) that the terms arising from γ are of the same order as those arising from λ, and that, with the actual values of these constants for the hydrogen halides, the former terms predominate for the outer lines of the overtone bands. It is probably therefore not justified to neglect them and retain λ.

* Schrödinger, *Ann. d. Phys.* vol. LXXIX, p. 484 (1926). See particularly Eq. (51).
† Mensing, *Zeit. f. Physik*, vol. XXXVI, p. 814 (1926).
‡ Born und Jordan, *Zeit. f. Physik*, vol. XXXIV, p. 858 (1925).

Digitized by Google

Digitized by Google

Digitized by Google

Digitized by Google

PROCEEDINGS

OF THE

Cambridge Philosophical Society.

The elliptic products of Jacobi and the theory of linear congruences. By Major P. A. MacMahon, St John's College.

[*Received* 22 May, *read* 26 July 1926.]

In the application of Elliptic Functions to the Theory of Numbers the two formulae of Jacobi

$$\prod_1^\infty (1-q^{2ma-a-b})(1-q^{2ma-a+b})(1-q^{2ma}) = \sum_{-\infty}^{+\infty} (-)^m q^{am^2+bm},$$

$$\prod_1^\infty (1+q^{2ma-a-b})(1+q^{2ma-a+b})(1-q^{2ma}) = \sum_{-\infty}^{+\infty} q^{am^2+bm}$$

are of great importance.

I denote either side of the first identity by

$$(a, b)$$

and either side of the second by

$$[a, b].$$

For my present purpose I take the exponents of q which occur on the left-hand sides to be integers. The summations are in regard to m, and a, b must be simultaneously either both integers or both the halves of uneven integers. If a, b possess a common factor f so that $a = a'f$, $b = b'f$, (a, b) and $[a, b]$ can be derived from (a', b'), $[a', b']$ respectively by writing q^f for q. It is then convenient to write

$$(a, b) = (a', b')_f; \quad [a, b] = [a', b']_f,$$

the subscript f implying that q^f is to be written for q.

With this understanding we may regard a, b as being relatively prime integers. Since moreover

$$(a, b) = -q^{a-b}(a, 2a-b) = -q^{a+b}(a, 2a+b) = (a, -b),$$

$$[a, b] = \quad q^{a-b}[a, 2a-b] = \quad q^{a+b}[a, 2a+b] = [a, -b],$$

we may always transform (a, b) so that $b \leqslant a$ and $[a, b]$ so that $b \leqslant a$.

In the first identity $(a, a) = 0$,

In the second $[a, a] = [\frac{1}{2}, \frac{1}{2}]_{2a} = 2[2, 1]_{2a}$.

Therefore in (a, b), for a given value of a, b may be any integer $< a$ and prime to a. It may therefore assume $\phi(a)$ values where ϕ is the ϕ function of Euler and $\phi(a)$ has been termed the totient of a by Sylvester and b one of the $\phi(a)$ totitives of a. If we have

Digitized by Google

before us $\left(\frac{a}{2}, \frac{b}{2}\right)$, where a, b are uneven integers, we may similarly regard b as one of the totitives of a. The same remarks apply to $[a, b]$ and $\left[\frac{a}{2}, \frac{b}{2}\right]$ in the two cases except that here b may be equal to a and then we have

$$[a, a] = [1, 1]_a, \quad \left[\frac{a}{2}, \frac{a}{2}\right] = \left[\frac{1}{2}, \frac{1}{2}\right]_a,$$

so that in this case the difference from the forms (a, b), $\left(\frac{a}{2}, \frac{b}{2}\right)$ is that those are the exceptional cases $[1, 1]$, $[\frac{1}{2}, \frac{1}{2}]$.

In recent years G. H. Hardy, S. Ramanujan and H. B. C. Darling have been concerned with those enumerating generating functions which involve elliptic products. They have obtained valuable theorems in regard to that part of the function which is entirely expressed by powers of q whose exponents $\equiv \epsilon \bmod \mu$ for special values of μ and ϵ. As a contribution to this study I investigate some properties of elliptic products which depend upon the partitioning of the product into parts each of which is specified by a certain congruence which is satisfied by the q exponents. In this partition the modulus μ remains constant and the residue varies from part to part. The residue cannot be assigned at pleasure, because it must be of the form $a\epsilon^2 + b\epsilon$, where ϵ is an integer.

§ 1. *The elliptic product* (a, b).

We put $\mu m + \epsilon$ for m in the series given above and obtain

$$(-)^{\epsilon} q^{a\epsilon^2 + b\epsilon} \sum_{-\infty}^{+\infty} (-)^{\mu m} q^{\mu^2 a m^2 + \mu(2a\epsilon + b)m}.$$

For the present I consider μ to be an uneven prime. We can then write the portion of (a, b), which has m of the form $\mu m + \epsilon$,

$$(-)^{\epsilon} q^{a\epsilon^2 + b\epsilon} (\mu a, 2a\epsilon + b)_{\mu}.$$

If a, b be both integers, every exponent of q in this function

$$\equiv a\epsilon^2 + b\epsilon \bmod \mu.$$

Moreover if a, b be the halves of uneven integers

$$a = \alpha + \tfrac{1}{2}, \quad b = \beta + \tfrac{1}{2},$$

the power of q is in general

$$(\alpha + \tfrac{1}{2})\epsilon^2 + (\beta + \tfrac{1}{2})\epsilon + \mu^2(\alpha + \tfrac{1}{2})m^2 + \mu(2a\epsilon + \epsilon + \beta + \tfrac{1}{2})m,$$

which is of the form $a\epsilon^2 + b\epsilon \bmod \mu$,

and $a\epsilon^2 + b\epsilon$ is an integer *provided* that

$$\binom{\mu m + 1}{2} \equiv 0 \bmod \mu;$$

Digitized by Google

and this is so because μ is uneven. Hence the transformed series is also

$$(-)^\epsilon q^{a\epsilon^2+b\epsilon}(\mu a, 2a\epsilon+b)_\mu,$$

where a, b are the halves of uneven integers.

We now sum this expression, giving ϵ the values $0, 1, 2, \ldots \mu-1$, and apply the formula

$$(a, b) = -q^{a-b}(a, 2a-b),$$

so as to make the second element in the bracket less than the first element, obtaining the standard formula

$$\begin{aligned}(a, b) = &-q^{a-b}(\mu a, 2a-b)_\mu + q^{4a-2b}(\mu a, 4a-b)_\mu - \ldots \\ &+(-)^{\frac{1}{2}(\mu-1)} q^{\left(\frac{\mu-1}{2}\right)^2 a-\frac{\mu-1}{2}b}(\mu a, \mu a-a-b)_\mu, \\ +(\mu a, b)_\mu &- q^{a+b}(\mu a, 2a+b)_\mu + q^{4a+2b}(\mu a, 4a+b)_\mu - \ldots \\ &+(-)^{\frac{1}{2}(\mu-1)} q^{\left(\frac{\mu-1}{2}\right)^2 a+\frac{\mu-1}{2}b}(\mu a, \mu a-a+b)_\mu,\end{aligned}$$

a partition into μ parts.

This formula is valid when a, b are integers or the halves of integers and is so far a complete solution of the question.

The partition is usually effective with less than μ parts.

If μ be even, a, b integers, we follow the foregoing method and find that the part separated is

$$(-)^\epsilon q^{a\epsilon^2+b\epsilon}[\mu a, 2a\epsilon+b]_\mu,$$

and it is to be observed that here the product (a, b) gives rise to products of type $[a, b]$.

Every exponent of q is clearly $\equiv a\epsilon^2+b\epsilon \bmod \mu$ and we obtain the crude partition

$$(a, b) = \sum_0^{\mu-1} (-)^\epsilon q^{a\epsilon^2+b\epsilon}[\mu a, 2a\epsilon+b]_\mu,$$

which by reason of the relation

$$[a, b] = q^{a-b}[a, 2a-b]$$

is reducible at once to the standard formula.

The elliptic product $[a, b]$.

The separated part is

$$q^{a\epsilon^2+b\epsilon}[\mu a, 2a\epsilon+b]_\mu.$$

If μ be even or uneven and a, b integers *or* if μ be uneven and a, b the halves of uneven integers, the standard formula is

Digitized by Google

$$[a, b] = + q^{a-b}[\mu a, 2a-b]_\mu + q^{4a-2b}[\mu a, 4a-b]_\mu + \ldots + q^{\left(\frac{\mu-1}{2}\right)^2 a - \frac{\mu-1}{2} b}[\mu a, \mu a - a - b]_\mu,$$
$$+ [\mu a, b]_\mu + q^{a+b}[\mu a, 2a+b]_\mu + q^{4a+2b}[\mu a, 4a+b]_\mu + \ldots + q^{\left(\frac{\mu-1}{2}\right)^2 a + \frac{\mu-1}{2} b}[\mu a, \mu a - a + b]_\mu.$$

$$(a, b) = - q^{a-b}[\mu a, 2a-b]_\mu + q^{4a-2b}[\mu a, 4a-b]_\mu - \ldots + (-)^{\frac{1}{2}(\mu-1)} q^{\left(\frac{\mu-1}{2}\right)^2 a - \frac{\mu-1}{2} b}[\mu a, \mu a - a - b]_\mu,$$
$$+ [\mu a, b]_\mu - q^{a+b}[\mu a, 2a+b]_\mu + q^{4a+2b}[\mu a, 4a+b]_\mu - \ldots + (-)^{\frac{1}{2}(\mu-1)} q^{\left(\frac{\mu-1}{2}\right)^2 + \frac{\mu-1}{2} b}[\mu a, \mu a - a + b]_\mu.$$

When μ is even, and a, b the halves of integers, the formula is not so simple or interesting but is readily obtainable.

A leading property of these partitions, which will be dealt with later, is that, when a is prime to μ, in one and in only one of the μ terms involved in the partition is the second element, which is of form

$$2ha \pm b,$$

divisible by μ. In this *special* term the subscript can be made equal to μ^2.

§ 2. *Important particular cases.*

The elliptic products which are of most importance from the present point of view are

$$(1, 0) = \prod_1^\infty (1-q^{2m-1})^2 (1-q^{2m}) = \sum_{-\infty}^{+\infty} (-)^m q^{m^2} = \sqrt{\frac{2k'K}{\pi}} = \theta_{0,1}(0),$$

$$[1, 0] = \prod_1^\infty (1+q^{2m-1})^2 (1-q^{2m}) = \sum_{-\infty}^{+\infty} q^{m^2} = \sqrt{\frac{2K}{\pi}} = \theta_{0,0}(0),$$

$$q^{\frac{1}{4}}[1, 1] = q^{\frac{1}{4}} \prod_1^\infty (1+q^{2m-2})(1-q^{4m}) = q^{\frac{1}{4}} \sum_{-\infty}^{+\infty} q^{m^2+m} = \sqrt{\frac{2kK}{\pi}} = \theta_{1,0}(0),$$

$$\left(\tfrac{3}{2}, \tfrac{1}{2}\right) = \prod_1^\infty (1-q^m) = \sum_{-}^{+\infty} (-)^m q^{\frac{3}{2}m^2 + \frac{1}{2}m}$$

(Euler's series),

Digitized by Google

$$(\tfrac{5}{2}, \tfrac{3}{2}) = \prod_1^\infty (1-q^{5m-4})(1-q^{5m-1})(1-q^{5m}) = \sum_{-\infty}^{+\infty} (-)^m q^{\frac{5}{2}m^2+\frac{3}{2}m},$$

$$(\tfrac{5}{2}, \tfrac{1}{2}) = \prod_1^\infty (1-q^{5m-3})(1-q^{5m-2})(1-q^{5m}) = \sum_{-\infty}^{+\infty} (-)^m q^{\frac{5}{2}m^2+\frac{1}{2}m}.$$

These last two arise in the Theory of Partitions by reason of properties that have been established in respect of the functions

$$\frac{(\frac{3}{2}, \frac{1}{2})_5}{(\frac{5}{2}, \frac{3}{2})}, \quad \frac{(\frac{3}{2}, \frac{1}{2})_5}{(\frac{5}{2}, \frac{1}{2})}$$

by L. Rogers* and S. Ramanujan†. They also arise in the elliptic function solutions of the general quintic equation as presented by H. B. C. Darling‡.

There is also the case of the Cube of Euler's series, the formula for which can be deduced from the fundamental formulae of Jacobi as a limiting case§.

The moduli of principal importance are prime numbers; in particular the early primes 2, 3, 5, 7, 11, ...

Postponing the case of the modulus 2, and generally of composite moduli, I proceed to the consideration of the simplest uneven prime moduli.

The modulus 3.

The standard formula is

$$(a, b) = (3a, b)_3 - q^{a+b}(3a, 2a+b)_3 - q^{a-b}(3a, 2a-b)_3,$$

$$[a, b] = [3a, b]_3 + q^{a+b}[3a, 2a+b]_3 + q^{a-b}[3a, 2a-b]_3.$$

Thence $$(1, 0) = (1, 0)_3 - 2q(3, 2)_3.$$

The special term establishes that

$$(1, 0) + 2q(3, 2)_3$$

is a transformation of the q series (1, 0), which is $\sqrt{\frac{2k'K}{\pi}}$, of order 3^2.

$$[1, 0] = [1, 0]_3 + 2q[3, 2]_3,$$

$$[1, 1] = 2[3, 1]_3 + q^2[1, 1]_3$$

indicating transformation of order 3^2 of

$$[1, 0] \text{ which is } \sqrt{\frac{2K}{\pi}},$$

* On Two Theorems, *Proc. L.M.S.*, New Series, vol. xvi, p. 315 *et seq.*

† Some properties of $p(n)$, the number of partitions of n, *Camb. Phil. Soc. Proc.*, vol. xix, p. 207 *et seq.*

‡ On The Trinomial Quintic, *Proc. L.M.S.*, New Series, vol. xxiii, p. 383 *et seq.*

§ Sylvester, *Collected Papers*, vol. iv, p. 60.

Digitized by Google

and $[1, 1]$ which is $q^{-\frac{1}{4}}\sqrt{\frac{2kK}{\pi}}$;

$$(\tfrac{3}{2}, \tfrac{1}{2}) = (\tfrac{9}{2}, \tfrac{1}{2})_3 - q\,(\tfrac{9}{2}, \tfrac{5}{2})_3 - q^2\,(\tfrac{9}{2}, \tfrac{7}{2})_3,$$

which, writing with Cayley $1 - q^s = (s)$, is

$$\begin{aligned}(1)\,(2)\,(3)\,\ldots &= (12)\,(39)\,(66)\ldots(15)\,(42)\,(69)\ldots(27)\,(54)\,(81)\ldots\\ &\quad - q\,(6)\,(33)\,(60)\ldots(21)\,(48)\,(75)\ldots(27)\,(54)\,(81)\ldots\\ &\quad - q^2\,(3)\,(30)\,(57)\ldots(24)\,(51)\,(78)\ldots(27)\,(54)\,(81)\ldots\end{aligned}$$

verifying a known result.

Since the leading element here is the half of an uneven number, viz. 3, which is not prime to the modulus 3, which we have before us, no special term arises.

The next two partitions are of particular interest and furnish the clue to an interesting theory

$$(\tfrac{5}{2}, \tfrac{3}{2}) = (\tfrac{5}{2}, \tfrac{1}{2})_{3^2} - q\,(\tfrac{15}{2}, \tfrac{7}{2})_3 - q\,\{q\,(\tfrac{15}{2}, \tfrac{13}{2})\}_3,$$

$$(\tfrac{5}{2}, \tfrac{1}{2}) = (\tfrac{15}{2}, \tfrac{1}{2})_3 - \{q\,(\tfrac{15}{2}, \tfrac{11}{2})\}_3 - q^2\,(\tfrac{5}{2}, \tfrac{3}{2})_{3^2}.$$

We find that the special term in $(\frac{5}{2}, \frac{3}{2})$ is $(\frac{5}{2}, \frac{1}{2})_{3^2}$,

" " " " $(\frac{5}{2}, \frac{1}{2})$ is $-q^2\,(\frac{5}{2}, \frac{3}{2})_{3^2}$.

This fact is remarkable, but not surprising, because the work to which allusion has been made has shewn that the two functions are very intimately connected not only in the Theory of the Transformation of Elliptic Functions but also in the application of Elliptic Functions to the solution of the general algebraic quintic equation. The property reminds one of that possessed by the so-called 'amicable numbers.' In this parallel the functions $(1, 0)$, $[1, 0]$, $[1, 1]$ considered above will be in correspondence with 'perfect numbers.'

We observe in the partitions of $(\frac{5}{2}, \frac{3}{2})$ and $(\frac{5}{2}, \frac{1}{2})$ that in the former case the special term is one part, the remaining two terms both belonging to the second part—whilst in the latter the special term is also one part and the remaining two terms constitute the second part.

Next we are led to consider the three products

$$(\tfrac{7}{2}, \tfrac{5}{2}),\ (\tfrac{7}{2}, \tfrac{3}{2}),\ (\tfrac{7}{2}, \tfrac{1}{2}),$$

a nd we find

$$(\tfrac{7}{2}, \tfrac{5}{2}) = (\tfrac{21}{2}, \tfrac{5}{2})_3 - \{q^2\,(\tfrac{21}{2}, \tfrac{19}{2})\}_3 - q\,(\tfrac{7}{2}, \tfrac{3}{2})_{3^2},$$

$$(\tfrac{7}{2}, \tfrac{3}{2}) = (\tfrac{7}{2}, \tfrac{1}{2})_{3^2} - q^2\,(\tfrac{21}{2}, \tfrac{11}{2})_3 - q^2\,\{q\,(\tfrac{21}{2}, \tfrac{17}{2})\}_3,$$

$$(\tfrac{7}{2}, \tfrac{1}{2}) = (\tfrac{21}{2}, \tfrac{1}{2})_3 - \{q\,(\tfrac{21}{2}, \tfrac{13}{2})\}_3 - q^4\,(\tfrac{7}{2}, \tfrac{5}{2})_{3^2}$$

Digitized by Google

exhibiting an amicable triad, for

$(\frac{7}{2}, \frac{5}{2})$ involves a definite *complete* part $-q\ (\frac{7}{2}, \frac{3}{2})_{3^2}$,

$(\frac{7}{2}, \frac{3}{2})$ „ „ „ „ $(\frac{7}{2}, \frac{1}{2})_{3^2}$,

$(\frac{7}{2}, \frac{1}{2})$ „ „ „ „ $-q^4\,(\frac{7}{2}, \frac{5}{2})_{3^2}$.

We have clearly before us a cyclic property.

So far we may say that it appears in the case of

$$\left.\begin{matrix}(1, 0)\\ [1, 0]\\ [1, 1]\\ (\frac{5}{2}, \frac{3}{2})\\ (\frac{5}{2}, \frac{1}{2})\end{matrix}\right\}\ \text{as (31)},$$

$$\left.\begin{matrix}(\frac{7}{2}, \frac{5}{2})\\ (\frac{7}{2}, \frac{3}{2})\\ (\frac{7}{2}, \frac{1}{2})\end{matrix}\right\}\ \text{as (531)},$$

the modulus being 3.

The modulus 5.

The standard partitions are

$$(a, b) = (5a, b)_5 - q^{a+b}\,(5a, 2a+b)_5 + q^{4a+2b}\,(5a, 4a+b)_5 - q^{a-b}\,(5a, 2a-b)_5 + q^{4a-2b}\,(5a, 4a-b)_5,$$

$$[a, b] = [5a, b]_5 + q^{a+b}\,[5a, 2a+b]_5 + q^{4a+2b}\,[5a, 4a+b]_5 + q^{a-b}\,[5a, 2a-b]_5 + q^{4a-2b}\,[5a, 4a-b]_5.$$

We derive

$$(1, 0) = (1, 0)_{5^2} - 2q\,(5, 2)_5 + 2q^4\,(5, 4)_5 \text{ three parts},$$

$$[1, 0] = [1, 0]_{5^2} + 2q\,[5, 2]_5 + 2q^4\,[5, 4]_5 \text{ three parts},$$

$$[1, 1] = 2\,[5, 1]_5 + 2q^2\,[5, 3]_5 + q^6\,[1, 1]_{5^2},$$

$$(\tfrac{3}{2}, \tfrac{1}{2}) = (\tfrac{15}{2}, \tfrac{1}{2})_5 + \{q\,(\tfrac{15}{2}, \tfrac{11}{2})\}_5 - q\,(\tfrac{3}{2}, \tfrac{1}{2})_{5^2},$$

$$- q^2\,(\tfrac{15}{2}, \tfrac{7}{2})_5 + q^2\,\{q\,(\tfrac{15}{2}, \tfrac{13}{2})\}_5.$$

In the partition, last written, the parts are given in separate rows so as to bring out the better that the special term stands out. It will be noticed that 3, the uneven integer of $(\frac{3}{2}, \frac{1}{2})$, being prime to the modulus, the cyclic property presents itself.

$$(\tfrac{5}{2}, \tfrac{3}{2}) = (\tfrac{25}{2}, \tfrac{3}{2})_5 - q^2\,(\tfrac{25}{2}, \tfrac{7}{2})_5 - q^4\,(\tfrac{25}{2}, \tfrac{13}{2})_5 + q^7\,(\tfrac{25}{2}, \tfrac{17}{2})_5 + q^{13}\,(\tfrac{25}{2}, \tfrac{23}{2})_5,$$

$$(\tfrac{5}{2}, \tfrac{1}{2}) = (\tfrac{25}{2}, \tfrac{1}{2})_5 - q^2\,(\tfrac{25}{2}, \tfrac{9}{2})_5 - q^3\,(\tfrac{25}{2}, \tfrac{11}{2})_5 + q^9\,(\tfrac{25}{2}, \tfrac{19}{2})_5 + q^{11}\,(\tfrac{25}{2}, \tfrac{21}{2})_5.$$

Digitized by Google

There is no cyclic property in evidence, as we expected, because the leading element 5 is not prime to the modulus. Each partition exhibits the full number of parts, viz. 5.

Next we have the triad

$$\begin{aligned}(\tfrac{7}{2}, \tfrac{5}{2}) &= (\tfrac{7}{2}, \tfrac{1}{2})_5 \\ &\quad - q\,(\tfrac{3}{2}\tfrac{5}{2}, \tfrac{9}{2})_5 - q^6\,(\tfrac{3}{2}\tfrac{5}{2}, \tfrac{19}{2})_5 \\ &\quad + q^9\,(\tfrac{3}{2}\tfrac{5}{2}, \tfrac{23}{2})_5 + q^{19}\,(\tfrac{3}{2}\tfrac{5}{2}, \tfrac{33}{2})_5, \\ (\tfrac{7}{2}, \tfrac{3}{2}) &= (\tfrac{3}{2}\tfrac{5}{2}, \tfrac{3}{2})_5 - q^5\,(\tfrac{3}{2}\tfrac{5}{2}, \tfrac{17}{2})_5 \\ &\quad + q^{11}\,(\tfrac{7}{2}, \tfrac{5}{2})_5 \\ &\quad - q^2\,(\tfrac{3}{2}\tfrac{5}{2}, \tfrac{11}{2})_5 + q^{17}\,(\tfrac{3}{2}\tfrac{5}{2}, \tfrac{31}{2})_5, \\ (\tfrac{7}{2}, \tfrac{1}{2}) &= (\tfrac{3}{2}\tfrac{5}{2}, \tfrac{1}{2})_5 + q^{15}\,(\tfrac{3}{2}\tfrac{5}{2}, \tfrac{29}{2})_5 \\ &\quad - q^3\,(\tfrac{3}{2}\tfrac{5}{2}, \tfrac{13}{2})_5 + q^{13}\,(\tfrac{3}{2}\tfrac{5}{2}, \tfrac{27}{2})_5 \\ &\quad - q^4\,(\tfrac{7}{2}, \tfrac{3}{2})_5,\end{aligned}$$

with the cyclic specification

(513) (cf. mod = 3).

The modulus 7.

The standard partitions are

$$\begin{aligned}(a, b) &= (7a, b)_7 \\ &\quad - q^{a+b}\,(7a, 2a+b)_7 + q^{4a+2b}\,(7a, 4a+b)_7 - q^{9a+3b}\,(7a, 6a+b)_7 \\ &\quad - q^{a-b}\,(7a, 2a-b)_7 + q^{4a-2b}\,(7a, 4a-b)_7 - q^{9a-3b}\,(7a, 6a-b)_7, \\ [a, b] &= [7a, b]_7 \\ &\quad + q^{a+b}\,[7a, 2a+b]_7 + q^{4a+2b}\,[7a, 4a+b]_7 + q^{9a+3b}\,[7a, 6a+b]_7 \\ &\quad + q^{a-b}\,[7a, 2a-b]_7 + q^{4a-2b}\,[7a, 4a-b]_7 + q^{9a-3b}\,[7a, 6a-b]_7, \\ (1, 0) &= (1, 0)_7 - 2q\,(7, 2)_7 + 2q^4\,(7, 4)_7 - 2q^6\,(7, 6)_7, \\ [1, 0] &= [1, 0]_7 + 2q\,[7, 2]_7 + 2q^4\,[7, 4]_7 + 2q^6\,[7, 6]_7, \\ [1, 1] &= 2\,[7, 1]_7 + 2q^2\,[7, 3]_7 + 2q^6\,[7, 5]_7 + q^{12}\,[1, 1]_7, \\ (\tfrac{3}{2}, \tfrac{1}{2}) &= (\tfrac{21}{2}, \tfrac{1}{2})_7 + \{q\,(\tfrac{21}{2}, \tfrac{19}{2})\}_7 \\ &\quad - q\,(\tfrac{21}{2}, \tfrac{5}{2})_7 - q\,\{q^2\,(\tfrac{21}{2}, \tfrac{19}{2})\}_7 \\ &\quad - q^2\,(\tfrac{3}{2}, \tfrac{1}{2})_7 \\ &\quad + q^5\,(\tfrac{21}{2}, \tfrac{11}{2})_7 - q^5\,\{q\,(\tfrac{21}{2}, \tfrac{17}{2})\}_7,\end{aligned}$$

having the cyclic specification (1).

Digitized by Google

$$
\begin{aligned}
(\tfrac{5}{2}, \tfrac{3}{2}) = \quad & (\tfrac{35}{2}, \tfrac{3}{2})_7 \quad + \{q\,(\tfrac{35}{2}, \tfrac{17}{2})\}_7 \\
& \qquad\qquad - q\,(\tfrac{5}{2}, \tfrac{1}{2})_{7^2} \\
& - q^4\,(\tfrac{35}{2}, \tfrac{13}{2})_7 - q^4\,\{q^2\,(\tfrac{35}{2}, \tfrac{27}{2})\}_7 \\
& + q^6\,(\tfrac{35}{2}, \tfrac{23}{2})_7 - q^6\,\{q^2\,(\tfrac{35}{2}, \tfrac{33}{2})\}_7, \\
(\tfrac{5}{2}, \tfrac{1}{2}) = \quad & (\tfrac{35}{2}, \tfrac{1}{2})_7 - \{q^3\,(\tfrac{35}{2}, \tfrac{29}{2})\}_7, \\
& - q^2\,(\tfrac{35}{2}, \tfrac{9}{2})_7 + q^2\,\{q\,(\tfrac{35}{2}, \tfrac{19}{2})\}_7 \\
& - q^3\,(\tfrac{35}{2}, \tfrac{11}{2})_7 - q^3\,\{q^2\,(\tfrac{35}{2}, \tfrac{31}{2})\}_7 \\
& + q^{11}\,(\tfrac{5}{2}, \tfrac{3}{2})_{7^2},
\end{aligned}
$$

a group of 2 having the cyclic specification (31).

The grouping of elliptic products is concerned with a product (a, b) or $[a, b]$ and a modulus μ such that μ and a are relatively prime and b one of the $\phi(a)$ totitives of a.

For the present I take μ to be a prime number so that the condition for a is that it is not a multiple of μ.

Before proceeding to its study as a branch of the Theory of Congruences in Higher Arithmetic it may be stated that the groups that will be considered or other groups have properties connected with the external powers of q which present themselves in the partitions.

The explanation of the Tables will be found in § 3.

The application to the partitions of elliptic products according to a given modulus may be instanced by reference to modulus 3 and leading element 14 of Table I.

The six totitives of the integer 14, viz. 13, 11, 9, 5, 3, 1, are given in two cycles of order three, viz.

$$(13, 5, 11)\ (1, 9, 3),$$

this means that the elliptic product (14, 13), when partitioned according to mod 3, has a part to a factor *près* $(14, 5)_{5^2}$, that (14, 5) has a part $(14, 11)_{5^2}$, and that (14, 11) has a part $(14, 13)_{5^2}$. Also that (14, 1), (14, 9), (14, 3) have parts $(14, 9)_{5^2}$, $(14, 3)_{5^2}$, $(14, 1)_{5^2}$ respectively.

When the whole of the totitives present themselves in a single cycle, I term the leading element a an elliptic primitive of the modulus μ.

The Table II is derived from Table I, obviously, by restriction to uneven integers in both elements of the elliptic product. When the whole of the uneven totitives of the uneven leading element numerator a present themselves in a single cycle we may, provisionally, term the leading element an uneven primitive of the modulus.

Digitized by Google

Table I.

a, b integers; b a totitive of a.

Mod 3.

First element a	Cycles of $\phi(a)$ numbers
2	(1)
4	(3, 1)
5	(4, 2) (1, 3)
7	(6, 2, 4) (1, 5, 3)
8	(7, 3, 1, 5)
10	(9, 3, 1, 7)
11	(10, 4, 2, 6, 8) (1, 7, 9, 5, 3)
13	(11, 5, 7) (1, 9, 3) (12, 4, 10) (2, 8, 6)
14	(13, 5, 11) (1, 9, 3)
16	(15, 5, 9, 3, 1, 11, 7, 13)
17	(16, 6, 2, 12, 4, 10, 8, 14) (1, 11, 15, 5, 13, 7, 9, 3)
19	(18, 6, 2, 12, 4, 14, 8, 10, 16) (1, 13, 17, 7, 15, 5, 11, 9, 3)

Primitives to the modulus 3

2, 4, 8, 10, 16, ...

Mod 5.

First element a	Cycles of $\phi(a)$ numbers
2	(1)
3	(2) (1)
4	(3, 1)
6	(5, 1)
7	(6, 4, 2) (1, 3, 5)
8	(7, 5, 1, 3)
9	(8, 2, 4) (1, 7, 5)
11	(10, 2, 4, 8, 6) (1, 9, 7, 3, 5)
12	(11, 7) (1, 5)
13	(12, 8) (9, 7) (11, 3) (2, 10) (4, 6) (1, 5)
14	(13, 3, 5, 1, 11, 9)
16	(15, 3, 7, 5, 1, 13, 9, 11)
17	(16, 10, 2, 14, 4, 6, 8, 12) (1, 7, 15, 3, 13, 11, 9, 5)
18	(17, 11, 5, 1, 7, 13)
19	(18, 4, 16, 12, 10, 2, 8, 6, 14) (1, 15, 3, 7, 9, 17, 11, 13, 5)

Elliptic primitives to the modulus 5

2, 4, 6, 8, 14, 16, 18, ...

Digitized by Google

TABLE I (*cont.*)

Mod 7.

First element a	Cycles of $\phi(a)$ numbers
2	(1)
3	(2) (1)
4	(3) (1)
5	(4, 2) (1, 3)
6	(5, 1)
8	(7, 1) (5, 3)
9	(8, 4, 2) (1, 5, 7)
10	(9, 7, 1, 3)
11	(10, 8, 2, 6, 4) (1, 3, 9, 5, 7)
12	(11, 5) (1, 7)
13	(12, 2, 4, 8, 10, 6) (1, 11, 9, 5, 3, 7)
15	(14, 2, 4, 8) (1, 13, 11, 7)
16	(15, 7, 1, 9) (13, 11, 3, 5)
17	(16, 12, 8, 6, 4, 14, 2, 10) (1, 5, 9, 11, 13, 3, 15, 7)
18	(17, 13, 7, 1, 5, 11)
19	(18, 8, 12) (16, 14, 2) (10, 4, 6) (1, 11, 7) (3, 5, 17) (9, 15, 13)

Elliptic primitives to the modulus 7

2, 6, 10, 18, ...

Mod 11.

First element a	Cycles of $\phi(a)$ numbers
2	(1)
3	(2) (1)
4	(3, 1)
5	(4) (3) (2) (1)
6	(5) (1)
7	(6, 2, 4) (1, 5, 3)
8	(7, 5, 1, 3)
9	(8, 4, 2) (1, 5, 7)
10	(9, 1) (7, 3)
12	(11, 1) (7, 5)
13	(12, 6, 10, 8, 4, 2) (1, 7, 3, 5, 9, 11)
14	(13, 9, 11, 1, 5, 3)
15	(14, 4) (13, 7) (2, 8) (1, 11)
16	(15, 13, 7, 11, 1, 3, 9, 5)
17	(16, 14, 8, 10, 4, 12, 2, 6) (1, 3, 9, 7, 13, 5, 15, 11)
18	(17, 5, 7) (1, 13, 11)
19	(18, 12, 8) (16, 2, 14) (10, 6, 4) (1, 7, 11) (3, 17, 5) (9, 13, 15)

Elliptic primitives to the modulus 11

2, 4, 8, 14, 16, ...

Digitized by Google

TABLE I (*cont.*)

Mod 13.

First element a	Cycles of $\phi(a)$ numbers
2	(1)
3	(2) (1)
4	(3, 1)
5	(4, 2) (1, 3)
6	(5) (1)
7	(6) (5) (4) (3) (2) (1)
8	(7, 3, 1, 5)
9	(8, 2, 4) (1, 7, 5)
10	(9, 7, 1, 3)
11	(10, 6, 8, 4, 2) (1, 5, 3, 7, 9)
12	(11, 1) (7, 5)
14	(13, 1) (11, 3) (9, 5)
15	(14, 8, 4, 2) (1, 7, 11, 13)
16	(15, 11, 9, 13, 1, 5, 7, 3)
17	(16, 4) (14, 12) (10, 6) (8, 2) (9, 15) (7, 11) (3, 5) (1, 13)
18	(17, 7, 5) (1, 11, 13)
19	(18, 16, 10, 8, 14, 4, 12, 2, 6) (1, 3, 9, 11, 5, 15, 7, 17, 13)

Elliptic primitives to the modulus 13

2, 4, 8, 10, 16, ...

Mod 17.

First element a	Cycles of $\phi(a)$ numbers
2	(1)
3	(2) (1)
4	(3) (1)
5	(4, 2) (1, 3)
6	(5, 1)
7	(6, 2, 4) (1, 5, 3)
8	(7) (5) (3) (1)
9	(8) (7) (5) (4) (2) (1)
10	(9, 3, 1, 7)
11	(9, 7, 3, 5, 1) (2, 4, 8, 6, 10)
12	(11, 5) (7, 1)
13	(12, 10, 4) (8, 2, 6) (5, 11, 7) (1, 3, 9)
14	(13, 9, 11, 1, 5, 3)
15	(14, 8, 4, 2) (1, 7, 11, 13)
16	(15, 1) (13, 3) (11, 5) (9, 7)
18	(17, 1) (13, 5) (11, 7)
19	(18, 10, 14, 12, 6, 16, 8, 4, 2) (1, 9, 5, 7, 13, 3, 11, 15, 17)

Elliptic primitives to the modulus 17

2, 6, 10, 14, ...

Digitized by Google

TABLE I (*cont.*)

Mod 19.

First element a	Cycles of $\phi(a)$ numbers
2	(1)
3	(2) (1)
4	(3, 1)
5	(4) (3) (2) (1)
6	(5, 1)
7	(6, 4, 2) (1, 3, 5)
8	(7, 3, 1, 5)
9	(8) (7) (5) (4) (2) (1)
10	(9) (7) (3) (1)
11	(10, 4, 6, 2, 8) (1, 7, 5, 9, 3)
12	(11, 7) (1, 5)
13	(12, 2, 4, 8, 10, 6) (1, 11, 9, 5, 3, 7)
14	(13, 11, 5) (1, 3, 9)
15	(1, 11) (14, 4) (8, 2) (7, 13)
16	(15, 11, 9, 13, 1, 5, 7, 3)
17	(16, 8, 4, 2) (12, 6, 14, 10) (5, 11, 3, 7) (1, 9, 13, 15)
18	(17, 1) (13, 5) (11, 7)

Primitives to the modulus 19

2, 4, 6, 8, 16, ...

TABLE II.

a, b both uneven integers; b a totitive of a.

Mod 3.

First element $\frac{a}{2}$	Cycles of $\phi'(a)$ numbers
$\frac{5}{2}$	$(\frac{3}{2}, \frac{1}{2})$
$\frac{7}{2}$	$(\frac{5}{2}, \frac{3}{2}, \frac{1}{2})$
$\frac{11}{2}$	$(\frac{9}{2}, \frac{5}{2}, \frac{3}{2}, \frac{1}{2}, \frac{7}{2})$
$\frac{13}{2}$	$(\frac{11}{2}, \frac{5}{2}, \frac{7}{2})$ $(\frac{1}{2}, \frac{9}{2}, \frac{3}{2})$
$\frac{17}{2}$	$(\frac{1}{2}, \frac{11}{2}, \frac{15}{2}, \frac{5}{2}, \frac{13}{2}, \frac{7}{2}, \frac{9}{2}, \frac{3}{2})$
$\frac{19}{2}$	.... $(\frac{1}{2}, \frac{13}{2}, \frac{17}{2}, \frac{7}{2}, \frac{15}{2}, \frac{5}{2}, \frac{11}{2}, \frac{9}{2}, \frac{3}{2})$

Uneven elliptic primitives

5, 7, 11, 17, 19, ...

Mod 5.

First element $\frac{a}{2}$	Cycles of $\phi'(a)$ numbers
$\frac{3}{2}$	$(\frac{1}{2})$
$\frac{7}{2}$	$(\frac{1}{2}, \frac{3}{2}, \frac{5}{2})$
$\frac{9}{2}$	$(\frac{1}{2}, \frac{7}{2}, \frac{5}{2})$
$\frac{11}{2}$	$(\frac{1}{2}, \frac{9}{2}, \frac{7}{2}, \frac{3}{2}, \frac{5}{2})$
$\frac{13}{2}$	$(\frac{11}{2}, \frac{3}{2})$ $(\frac{9}{2}, \frac{7}{2})$ $(\frac{5}{2}, \frac{1}{2})$
$\frac{17}{2}$	$(\frac{15}{2}, \frac{3}{2}, \frac{13}{2}, \frac{11}{2}, \frac{9}{2}, \frac{5}{2}, \frac{1}{2}, \frac{7}{2})$
$\frac{19}{2}$	$(\frac{1}{2}, \frac{15}{2}, \frac{3}{2}, \frac{7}{2}, \frac{9}{2}, \frac{17}{2}, \frac{11}{2}, \frac{13}{2}, \frac{5}{2})$

Uneven elliptic primitives

3, 7, 9, 11, 17, 19, ...

Digitized by Google

TABLE II (*cont.*)

Mod 7.

First element $\frac{a}{2}$	Cycles of $\phi'(a)$ totitives
$\frac{1}{2}$	$(\frac{1}{2})$
$\frac{3}{2}$	$(\frac{3}{2}, \frac{1}{2})$
$\frac{5}{2}$	$(\frac{7}{2}, \frac{1}{2}, \frac{5}{2})$
$\frac{9}{2}$	$(\frac{9}{2}, \frac{5}{2}, \frac{7}{2}, \frac{1}{2}, \frac{3}{2})$
$\frac{11}{2}$	$(\frac{11}{2}, \frac{9}{2}, \frac{5}{2}, \frac{3}{2}, \frac{7}{2}, \frac{1}{2})$
$\frac{13}{2}$	$(\frac{13}{2}, \frac{11}{2}, \frac{7}{2}, \frac{1}{2})$
$\frac{15}{2}$	$(\frac{1}{2}, \frac{5}{2}, \frac{9}{2}, \frac{11}{2}, \frac{13}{2}, \frac{3}{2}, \frac{15}{2}, \frac{7}{2})$
$\frac{17}{2}$	$(\frac{1}{2}, \frac{11}{2}, \frac{7}{2})(\frac{3}{2}, \frac{5}{2}, \frac{17}{2})(\frac{9}{2}, \frac{15}{2}, \frac{13}{2})$
$\frac{19}{2}$	

Uneven elliptic primitives

3, 5, 9, 11, 13, 15, 17, ...

Mod 11.

First element $\frac{a}{2}$	Cycles of $\phi'(a)$ totitives
$\frac{1}{2}$	$(\frac{1}{2})$
$\frac{3}{2}$	$(\frac{3}{2})(\frac{1}{2})$
$\frac{5}{2}$	$(\frac{5}{2}, \frac{3}{2}, \frac{1}{2})$
$\frac{7}{2}$	$(\frac{7}{2}, \frac{1}{2}, \frac{5}{2})$
$\frac{9}{2}$	$(\frac{11}{2}, \frac{1}{2}, \frac{7}{2}, \frac{3}{2}, \frac{5}{2}, \frac{9}{2})$
$\frac{11}{2}$	$(\frac{13}{2}, \frac{7}{2})(\frac{11}{2}, \frac{1}{2})$
$\frac{13}{2}$	$(\frac{1}{2}, \frac{3}{2}, \frac{9}{2}, \frac{7}{2}, \frac{13}{2}, \frac{5}{2}, \frac{15}{2}, \frac{11}{2})$
$\frac{15}{2}$	$(\frac{1}{2}, \frac{7}{2}, \frac{11}{2})(\frac{3}{2}, \frac{17}{2}, \frac{5}{2})(\frac{9}{2}, \frac{13}{2}, \frac{15}{2})$
$\frac{17}{2}$	
$\frac{19}{2}$	

Uneven elliptic primitives

3, 7, 9, 13, 17, ...

Mod 13.

First element $\frac{a}{2}$	Cycles of $\phi'(a)$ totitives
$\frac{1}{2}$	$(\frac{1}{2})$
$\frac{3}{2}$	$(\frac{3}{2}, \frac{1}{2})$
$\frac{5}{2}$	$(\frac{5}{2})(\frac{3}{2})(\frac{1}{2})$
$\frac{7}{2}$	$(\frac{7}{2}, \frac{5}{2}, \frac{1}{2})$
$\frac{9}{2}$	$(\frac{9}{2}, \frac{1}{2}, \frac{5}{2}, \frac{3}{2}, \frac{7}{2})$
$\frac{11}{2}$	$(\frac{13}{2}, \frac{1}{2}, \frac{7}{2}, \frac{11}{2})$
$\frac{13}{2}$	$(\frac{9}{2}, \frac{15}{2})(\frac{7}{2}, \frac{11}{2})(\frac{3}{2}, \frac{5}{2})(\frac{1}{2}, \frac{13}{2})$
$\frac{15}{2}$	$(\frac{1}{2}, \frac{3}{2}, \frac{9}{2}, \frac{11}{2}, \frac{5}{2}, \frac{15}{2}, \frac{7}{2}, \frac{17}{2}, \frac{13}{2})$
$\frac{17}{2}$	
$\frac{19}{2}$	

Uneven elliptic primitives

3, 5, 9, 15, 19, ...

Digitized by Google

TABLE II (*cont.*)

Mod 17.

First element $\frac{a}{2}$	Cycles of $\phi'(a)$ totitives
$\frac{3}{2}$	$(\frac{1}{2})$
$\frac{5}{2}$	$(\frac{3}{2}, \frac{1}{2})$
$\frac{7}{2}$	$(\frac{5}{2}, \frac{3}{2}, \frac{1}{2})$
$\frac{9}{2}$	$(\frac{7}{2})(\frac{5}{2})$
$\frac{11}{2}$	$(\frac{9}{2}, \frac{7}{2}, \frac{3}{2}, \frac{5}{2}, \frac{1}{2})$
$\frac{13}{2}$	$(\frac{11}{2}, \frac{7}{2}, \frac{5}{2})(\frac{9}{2}, \frac{1}{2}, \frac{3}{2})$
$\frac{15}{2}$	$(\frac{13}{2}, \frac{11}{2})(\frac{7}{2}, \frac{1}{2})$
$\frac{19}{2}$	$(\frac{1}{2}, \frac{9}{2}, \frac{5}{2}, \frac{7}{2}, \frac{13}{2}, \frac{3}{2}, \frac{11}{2}, \frac{15}{2}, \frac{17}{2})$

Uneven elliptic primitives

3, 5, 7, 11, 19, ...

Mod 19.

First element $\frac{a}{2}$	Cycles of $\phi'(a)$ totitives
$\frac{3}{2}$	$(\frac{1}{2})$
$\frac{5}{2}$	$(\frac{3}{2})(\frac{1}{2})$
$\frac{7}{2}$	$(\frac{1}{2}, \frac{3}{2}, \frac{5}{2})$
$\frac{9}{2}$	$(\frac{7}{2})(\frac{5}{2})(\frac{1}{2})$
$\frac{11}{2}$	$(\frac{1}{2}, \frac{7}{2}, \frac{5}{2}, \frac{9}{2}, \frac{3}{2})$
$\frac{13}{2}$	$(\frac{1}{2}, \frac{11}{2}, \frac{9}{2}, \frac{5}{2}, \frac{3}{2}, \frac{7}{2})$
$\frac{15}{2}$	$(\frac{1}{2}, \frac{11}{2})(\frac{7}{2}, \frac{13}{2})$
$\frac{17}{2}$	$(\frac{5}{2}, \frac{11}{2}, \frac{3}{2}, \frac{7}{2})(\frac{1}{2}, \frac{9}{2}, \frac{13}{2}, \frac{15}{2})$

Uneven elliptic primitives

3, 7, 11, 13, ...

§ 3.

When we regard the annexed Tables certain properties appear to be in evidence which now come up for examination.

μ being an uneven prime, a any integer prime to μ and b any one of the $\phi(a)$ totitives of a, the integers b appear in the elliptic product discussion in two series each of which is in arithmetical progression, viz.

$b,\ 2a+b,\ 4a+b,\ \ldots (\mu-1)a+b \quad \frac{1}{2}(\mu+1)$ terms $\qquad (A)$,

$2a-b,\ 4a-b,\ \ldots (\mu-1)a-b \quad \frac{1}{2}(\mu-1)$ terms $\qquad (B)$.

Write $\qquad 2ka+b \equiv \alpha_k \bmod \mu,$

$\qquad 2ka-b \equiv \beta_k \bmod \mu,$

so as to obtain the two series

$\alpha_0,\ \alpha_1,\ \alpha_2,\ \ldots\ \alpha_{\frac{1}{2}(\mu-1)} \qquad (A')$,

$\beta_1,\ \beta_2,\ \ldots\ \beta_{\frac{1}{2}(\mu-1)} \qquad (B')$.

Digitized by Google

Since the common difference in each of the series (A), (B) is $2a$ and $2a$ is prime to μ, we know that the residues which appear in (A') are all different—so also the residues which appear in (B'). If $2a \equiv \delta \bmod \mu$, the residues in (A') and in (B') may be taken to have the common difference δ.

The series are thus

$$\alpha_0,\ \alpha_0+\delta,\ \alpha_0+2\delta,\ \ldots\ \alpha_0+\tfrac{1}{2}(\mu-1)\,\delta,$$
$$\beta_1,\quad \beta_1+\delta,\ \ldots\ \beta_1+\tfrac{1}{2}(\mu-3)\,\delta,$$

when reduced by the relations

$$\alpha_0+k\delta \equiv \alpha_k \bmod \mu,$$
$$\beta_1+(k-1)\,\delta \equiv \beta_k \bmod \mu.$$

If the series do not terminate with the terms $(\mu-1)\,a \pm b$ but are continued indefinitely, we write

$$b,\ 2a+b,\ 4a+b,\ \ldots\ (A_\infty),$$
$$2a-b,\ 4a-b,\ \ldots\ (B_\infty);$$
$$\alpha_0,\ \alpha_1,\ \alpha_2,\ \ldots\qquad (A'_\infty),$$
$$\beta_1,\ \beta_2,\ \ldots\qquad (B'_\infty).$$

By a known theorem the first μ terms in each of the series (A'_∞), (B'_∞) are the μ integers 0, 1, 2, ... $\mu-1$ in some order. Thus for some values of i and j

$$\alpha_i = \beta_j.$$

This being so it is clear that the two series

$$\alpha_i,\ \alpha_{i+1},\ \ldots\ \alpha_{i+\mu-1},$$
$$\beta_j,\ \beta_{j+1},\ \ldots\ \beta_{j+\mu-1}$$

are identical.

Theorem I. **Either** the series (A') involves the residue zero and (B') does not **or** (B') involves zero and (A') does not.

For suppose k and l to be integers such that

$$2ka+b \equiv 0 \bmod \mu,\quad 2la-b \equiv 0 \bmod \mu,$$

giving
$$2\,(k+l)\,a \equiv 0 \bmod \mu$$

and
$$k+l \equiv 0 \bmod \mu.$$

If $k+l=\mu$, we can only have $k=\frac{1}{2}(\mu+1)$, $l=\frac{1}{2}(\mu-1)$, an impossibility because we cannot have simultaneously

$$(\mu-1)\,a+b \equiv 0 \bmod \mu,$$
$$(\mu-1)\,a-b \equiv 0 \bmod \mu.$$

Hence zeros cannot occur in each of the series (A'), (B'). But the first μ terms in each of the series (A'_∞), (B'_∞) involves a zero, so

Digitized by Google

that we may have $k+l<2\mu$. Hence for some values of k, l, $k+l=\mu$. Hence one and one only of the series (A'), (B') involves zero.

In a similar manner it may be shewn that if the residue $+\epsilon$ occurs in the series (A') the residue $-\epsilon$ cannot occur in (B').

It follows that if b_1 be any one of the $\phi(a)$ totitives of a, we can uniquely select a term from the combined series (A), (B) which is congruent to zero for the modulus μ.

Dividing this term by μ we reach the number b_2 which is also a totitive of a; thence from b_2 we reach a totitive b_3 and so on until finally we reach a totitive b_{s+1} which is equal to b_1, and observe the cycle of totitives

$$(b_1, b_2, \ldots b_s) \text{ of order } s.$$

It thus appears that the whole of the totitives of a can be arranged in a set of cycles of various orders as exhibited in the table for given simple values of μ and a.

Theorem II. If in any cycle of totitives b' exceeds b, then in the same or some other cycle the totitive $a-b'$ succeeds the totitive $a-b$.

Let θ be a symbol such that $\theta^2=1$; θ may be therefore either $+1$ or -1. Since

$$2ia+\theta b=\mu b'$$

for some value of i, assume

$$2ja+\theta(a-b)=\mu(a-b')$$

for some value of j; by addition

$$2(i+j)a+\theta a=\mu a$$

or
$$2j=\mu-\theta-2i$$

and
$$j=\tfrac{1}{2}(\mu-\theta-2i)$$

uniquely. Hence the succession $a-b$, $a-b'$ follows from the succession b, b' and the theorem is proved.

In particular the succession b, b' implies the succession b', b, and there is a cycle of order 2, when

$$a=b+b',$$

$$2ia+\theta b=\mu(a-b),$$

$$(\mu-2i)a=(\mu+\theta)b,$$

giving
$$a=\mu+\theta,\quad b=\mu-2i,$$

so that b is uneven and $a=\mu\pm1$.

Digitized by Google

Theorem III. A cycle of order 2

$$(b, a-b)$$

occurs when $a = \mu \pm 1$, b an uneven number.

Let $\theta_t^2 = 1$ so that we may put $\theta_t = +1$ or -1 at pleasure.

In the case of a cycle of order s we have s equations

$$\begin{aligned} 2i_1 a + \quad \theta_1 b_1 \quad &= \mu b_2, \\ 2i_2 a + \quad \theta_2 b_2 \quad &= \mu b_3, \\ &\cdots\cdots\cdots\cdots \\ 2i_{s-1} a + \theta_{s-1} b_{s-1} &= \mu b_s, \\ 2i_s a + \quad \theta_s b_s \quad &= \mu b_1, \end{aligned}$$

and we find:

Theorem IV. The integers in the cycle

$$(b_1, b_2, \ldots b_s)$$

are all of the same parity. As a result of this we have:

Theorem V. If the order of the cycle be the maximum $\phi(a)$, a must be an even number.

Theorem VI. An elliptic primitive to the uneven prime modulus μ must be an even number.

Theorem VII. The cycles in which the totitives of a appear for a given modulus are all of the same order.

From the s equations above written we can eliminate the $s-1$ quantities $b_2, b_3, \ldots b_s$ and obtain the relation

$$2(i_s\mu^{s-1} + \theta_s i_{s-1}\mu^{s-2} + \theta_s\theta_{s-1} i_{s-2}\mu^{s-3} + \ldots + \theta_s\theta_{s-1}\ldots\theta_2 i_1)\, a = (\mu^s - \theta_1\theta_2\ldots\theta_s)\, b_1,$$

and by circular procession we can write down $s-1$ other relations of which the next is

$$2(i_1\mu^{s-1} + \theta_1 i_s\mu^{s-2} + \theta_1\theta_s i_{s-1}\mu^{s-3} + \ldots + \theta_1\theta_s\ldots\theta_3 i_2)\, a = (\mu^s - \theta_1\theta_2\ldots\theta_s)\, b_2,$$

and the ultimate one is

$$2(i_{s-1}\mu^{s-1} + \theta_{s-1} i_{s-2}\mu^{s-2} + \theta_{s-1}\theta_{s-2} i_{s-3}\mu^{s-3} + \ldots + \theta_1\theta_2\ldots\theta_{s-1} i_s)\, a = (\mu^s - \theta_1\theta_2\ldots\theta_s)\, b_s.$$

Any integer i_t may assume values

$$0, 1, 2, \ldots \tfrac{1}{2}(\mu - 1),$$

and θ_t may be $+1$ or -1, but we may not put $i_t = 0$, $\theta_t = -1$ simultaneously.

Digitized by Google

Adding the s equations

$$2(\mu^{s-1}\Sigma i_s + \mu^{s-2}\Sigma\theta_s i_{s-1} + \ldots + \Sigma\theta_s\theta_{s-1}\ldots\theta_2 i_1)\,a = (\mu^s - \theta_1\theta_2\ldots\theta_s)\,\Sigma b.$$

The product $\theta_1\theta_2\ldots\theta_s = \pm 1$, and the bracketed factors on each side of the tth of the s equations must possess a common factor so as to reduce their ratio to $\frac{2a}{b_t}$. Moreover, in the last equation written, the ratio of the factors must reduce to $\frac{2a}{\Sigma b}$.

I hope to resume consideration of this subject on a future occasion.

Digitized by Google

On Mellin's inversion formula. By Mr J. C. BURKILL, Trinity College.

[*Received* 29 May, *read* 26 July 1926.]

1. The extension of Mellin's inversion formula expressed by the equations

$$f(s) = \int_0^\infty y^s \, d\phi(y),$$

$$\phi(x) = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} f(s)\, x^s \frac{ds}{s}$$

has been considered by Fowler* who shows that some form of Stieltjes integral is essential to Poincaré's proof of the necessity of the quantum hypothesis. Fowler confines his discussion to a restricted type of function $\phi(y)$ which is sufficient for the physical problem. It will be proved here that the formulae hold with a general Stieltjes integral in the first equation.

I have remarked elsewhere† that inversion formulae in Stieltjes integrals arise from integrals of the "discontinuous factor" type. The discontinuous integral in the present instance is

$$\frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} k^s \frac{ds}{s} \qquad (k > 0),$$

which is equal to 1, $\frac{1}{2}$ or 0 according as $k >, =$ or < 1, the integral being interpreted as the principal value

$$\lim_{T \to \infty} I(k, T),$$

where

$$I(k, T) = \frac{1}{2\pi i} \int_{a-iT}^{a+iT} k^s \frac{ds}{s}.$$

2. We need two lemmas on the behaviour of $I(k, T)$.

Lemma 1. *If* $a > 0$, $T > 0$, *then*

$$|I(k, T) - 1| \leqslant \frac{1}{\pi T} \frac{k^a}{\log k} \qquad (k > 1),$$

$$|I(k, T)| \leqslant \frac{1}{\pi T} \frac{k^a}{\log (1/k)} \qquad (k < 1),$$

$$|I(k, T) - \tfrac{1}{2}| \leqslant \frac{a}{\pi T} \qquad (k = 1).$$

* *Proc. Royal Soc.* (A), vol. 99 (1921), pp. 462–471. An account of Mellin's formula in the ordinary form (without Stieltjes integrals) is given by Hardy, *Messenger of Math.* vol. 47 (1918), pp. 178–184 and vol. 50 (1921), pp. 165–171.

† "The expression in Stieltjes integrals of the inversion formulae of Fourier and Hankel," *Proc. London Math. Soc.* (unpublished).

Digitized by Google

Cor. $I(k, T)$ *tends to its limit uniformly in the intervals* $0 < \delta \leqslant k \leqslant 1 - \delta$ *and* $1 + \delta \leqslant k \leqslant k_0$.

For proofs, see Landau, *Primzahlen*, 342.

Lemma 2. *As* $T \to \infty$,

$$I(k, T) = O(1)$$

uniformly in k *for* $0 < k_0 \leqslant k \leqslant k_1$.

Writing $m = \log k$—so that m lies between fixed bounds—we have

$$\begin{aligned} I(k, T) &= \frac{k^a}{2\pi} \int_{-T}^{T} \frac{\cos mt + i \sin mt}{a + it}\, dt \\ &= \frac{k^a}{2\pi} \int_{-T}^{T} \frac{a \cos mt + t \sin mt}{a^2 + t^2}\, dt \\ &= O(1) + O\left| \int_0^{mT} \frac{u \sin u}{m^2 a^2 + u^2}\, du \right|. \end{aligned}$$

Since $u^2/(m^2a^2 + u^2)$ is an increasing function of u, the second mean value theorem shows that the integral in the second term is

$$\frac{T^2}{a^2 + T^2} \int_\alpha^{mT} \frac{\sin u}{u}\, du,$$

where α is between 0 and mT, which is $O(1)$, and the lemma is proved.

3. We can now prove the main theorem.

Theorem A. *If*

(1) $\phi(y)$ *has bounded variation in* $0 \leqslant y \leqslant k$ *for every* $k > 0$,

(2) $\phi(y) = \frac{1}{2}\{\phi(y + 0) + \phi(y - 0)\}$,

(3) $\phi(+0) = 0$,

(4) $\int_0 y^{-\alpha} |\, d\phi(y)\,|$ *converges where* $\alpha > 0$

and $\int^\infty y^{-\beta} |\, d\phi(y)\,|$ *converges where* $\beta < \alpha$,

and if

$$f(s) = \int_0^\infty y^{-s}\, d\phi(y) \qquad (\beta < \sigma < \alpha), \tag{3·1}$$

then will

$$\phi(x) = \frac{1}{2\pi i} \int_{a - i\infty}^{a + i\infty} f(s)\, x^s \frac{ds}{s} \quad (\beta < a < \alpha,\ a > 0), \tag{3·2}$$

where the integrals (3·1), (3·2) *are respectively* $\lim \int_\epsilon^\infty$, $\lim \int_{a - iT}^{a + iT}$.

Digitized by Google

From (4), $\int_0^\infty y^{-s}d\phi(y)$ is uniformly convergent in any rectangle

$$\alpha < \alpha' \leqslant \sigma \leqslant \beta' < \beta, \quad -T \leqslant t \leqslant T$$

and $f(s)$ is regular in $\alpha < \sigma < \beta$.

Substituting the value of $f(s)$ from (3·1) in the right-hand side of (3·2), we have

$$\lim_{T\to\infty} \frac{1}{2\pi i}\int_{a-iT}^{a+iT} \frac{x^s}{s}\,ds \int_0^\infty y^{-s}\,d\phi(y).$$

Since the inner integral is uniformly convergent for the values of s in question, we may invert the order of integration and obtain

$$\lim_{T\to\infty} \int_0^\infty I\left(\frac{x}{y}, T\right) d\phi(y).$$

Split up this integral into

$$\int_0^\eta + \int_\eta^{x-\delta} + \int_{x-\delta}^{x+\delta} + \int_{x+\delta}^{k} + \int_k^\infty.$$

By Lemma 1, the integrand in the first and last integrals is $O(y^{-a})$ and in the second and fourth it converges uniformly to its limit as $T\to\infty$; and in the third integral the integrand is bounded by Lemma 2.

The proof is completed by choice of η, k, δ in exactly the same way as that of the main theorem in my Fourier-Stieltjes paper referred to.

4. If in Theorem A we replace y by e^y we obtain the exponential form of Fourier's inversion formula.

Theorem B. *If*

(1) $\phi(y)$ *has bounded variation in* $0 \leqslant y \leqslant k$ *for every* $k > 0$,

(2) $\phi(y) = \frac{1}{2}\{\phi(y+0) + \phi(y-0)\}$,

(3) $\phi(0) = 0$,

(4) $\int^\infty e^{-\beta y}\,|\,d\phi(y)\,|$ *converges,*

and if

$$f(s) = \int_0^\infty e^{-sy}\,d\phi(y) \qquad (\beta < \sigma), \tag{4·1}$$

then will

$$\phi(x) = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} f(s)\,e^{sx}\,\frac{ds}{s} \qquad (\beta < a,\ a > 0), \tag{4·2}$$

where the integral (4·2) *is* $\lim \int_{a-iT}^{a+iT}$.

Digitized by Google

If in particular $\phi(y)$ is a step-function, this reduces to the Cahen-Hadamard-Perron formula for the sum of the first n coefficients of a Dirichlet series.

5. We may extend Theorem B to give an inversion formula involving Bessel functions of the third kind instead of exponentials by starting with an appropriate discontinuous factor. Such a factor is

$$(5{\cdot}1) \qquad \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} \frac{\pi}{2} H_\mu^{(1)}(isy)\, y^{-\mu} H_{\mu+1}^{(2)}(isx)\, x^{\mu+1}\, ds,$$

where $x>0$, $y>0$, $a>0$, which has the value 1, $\frac{1}{2}$ or 0 according as $y<$, $=$ or $>x$. This integral has been applied by Oppenheim* to the discussion of lattice-point problems.

We need lemmas corresponding to (1) and (2) of § 4. These are easily established by writing the integral (5·1) as

$$\frac{1}{2\pi i}\left\{\int_{a-iT}^{a+iT} + \int_{a-i\infty}^{a-iT} + \int_{a+iT}^{a+i\infty}\right\},$$

substituting in the last two integrals the asymptotic value of the integrand †

$$\frac{e^{s(x-y)}}{s}\left(\frac{x}{y}\right)^{\mu+\frac{1}{2}}\left\{1+O\left(\frac{1}{t}\right)\right\}$$

and then applying the results of lemmas (1) and (2) above with e^{x-y} in place of k.

In generalising from exponentials to Bessel functions an extra condition—(4) below—has to be introduced to secure the convergence of the Stieltjes integral at the origin. The precise statement is as follows.

Theorem C. *If*

(1) $\phi(y)$ *has bounded variation in* $0 \leqslant y \leqslant k$ *for every* $k>0$,

(2) $\phi(y) = \frac{1}{2}\{\phi(y+0)+\phi(y-0)\}$,

(3) $\phi(+0)=0$,

(4) $\int_0 \lambda(y)\,|\,d\phi(y)\,|$ *converges,*

where $\lambda(y)$ *is* $y^{-2\mu}$, $\log y$ *or* 1

according as $\mu>$, $=$ *or* <0

and $\int^\infty e^{-\beta y}\,|\,d\phi(y)\,|$ *converges,*

* A. Oppenheim, "Some identities in the theory of numbers," *Proc. London Math. Soc.* (*Records*), vol. 24 (1925), p. xxiii. I am indebted to Mr Oppenheim for sending me in MS. the part of his work dealing with discontinuous factors.

† Watson, *Bessel Functions*, p. 198.

Digitized by Google

and if

$$f(s) = \int_0^\infty \left(\frac{\pi}{2}\right)^{\frac{1}{2}} H_\mu^{(1)}(isy)\, y^{-\mu}\, d\phi(y) \qquad (\beta < \sigma), \tag{5·2}$$

then will

$$\phi(x) = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} \left(\frac{\pi}{2}\right)^{\frac{1}{2}} H_{\mu+1}^{(2)}(isx)\, x^{\mu+1} f(s)\, ds, \tag{5·3}$$

where $\beta < a$, $a > 0$ and the integral is the principal value.

From a different discontinuous factor we obtain the pair of formulae

$$f(s) = \int_0^\infty \left(\frac{\pi}{2}\right)^{\frac{1}{2}} H_\mu^{(1)}(isy)\, y^{\mu}\, d\phi(y), \tag{5·4}$$

$$\phi(x) = -\frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} \left(\frac{\pi}{2}\right)^{\frac{1}{2}} H_{\mu-1}^{(2)}(isx)\, x^{-\mu+1} f(s)\, ds, \tag{5·5}$$

provided that we now define $\lambda(y)$ in condition (4) to be

$$1,\ \log y \text{ or } y^{-2\mu}$$

according as $$\mu >, = \text{ or } < 0.$$

Digitized by Google

Note on the extension to higher space of a theorem of Wallace.
By Mr J. P. GABBATT, Peterhouse.

[*Received* 18 June, *read* 26 July 1926.]

The theorem is attributed to Wallace* that, in a euclidean plane, the circumcircles of the triangles determined by four lines, of general position, meet at a point. It is further known† that, in euclidean space of n dimensions, the circumhyperspheres of the simplices determined by $n+2$ flats, of general position, meet at a point, if and only if n be even.

The following closely related theorem is also known‡: In euclidean space of n dimensions (whether n be odd or even), if the $n+1$ points $A_{00}, A_{11}, \ldots, A_{nn}$ determine the base-simplex, and if A_{ij} denote any point on the edge $A_{ii}A_{jj}$ $(i, j = 0, 1, \ldots, n)$; then the $n+1$ hyperspheres, $[S_i]$, determined respectively by the $n+1$ groups of points $A_{i0}, A_{i1}, \ldots, A_{in}$ meet at a point. Haskell has remarked† that, if and only if the $n+1$ radical flats of the circumhypersphere of the base-simplex, and the $n+1$ hyperspheres $[S_i]$, respectively, meet at a point, then that point is common, not only to the hyperspheres $[S_i]$, but also to the circumhypersphere; and that the analytical expression of the appropriate condition for this, when the $\frac{1}{2}n(n+1)$ points A_{ij} lie on a flat, is the vanishing of a certain skew-symmetric determinant, Δ, of order $n+1$. If n be even, such a determinant vanishes in all cases; and the extension of Wallace's theorem to space of even dimensions follows.

If n be odd, the determinant Δ does not generally vanish, but is the square of a function, P, the Pfaffian of the elements of Δ; and P is a factor of all the (non-vanishing) first minors of Δ§. Thus the vanishing, as a special case, of Δ involves the vanishing of all its first minors. It does not seem to have been noticed that this amounts to the theorem that, if n be odd, and if the $n+1$ radical flats, referred to above, have a point common to all, then they have a line common to all; and therefore that: *In euclidean space of n dimensions, where n is odd; if $n+2$ given flats satisfy a single*

* 'Scoticus,' *Leybourn's Math. Repos.*, N.S., 1 (1806), 170; see Mackay, *Proc. Edin. Math. Soc.* 9 (1891), 87.

† Grace, *Trans. Camb. Phil. Soc.*, 16 (1898), 153–190 (163); Kühne, *Crelle*, 119 (1898), 186–195 (corrected by H. F. Baker, *Proc. Camb. Phil. Soc.*, 22 (1924), 28–33); Haskell, *Arch. d. Math. u. Phys.* (3), 5 (1903), 278–281.

‡ For two dimensions, Miquel, *Liouville*, 3 (1838), 485–487; for three dimensions, S. Roberts, *Proc. Lond. Math. Soc.*, 12 (1881), 102, 117–120, *ibid.* 25 (1894), 306–314; for four dimensions, Grace, *loc. cit.*, (168); for n dimensions, Haskell, *oc. cit.*, W. F. Meyer, *Arch. d. Math. u. Phys.* (3), 5 (1903), 282–287.

§ See *e.g.* Scott and Mathews, *Theory of Determinants*, ed. 2, Cambridge (1904), 92–96.

Digitized by Google

condition, then the circumhyperspheres of the $n+2$ simplices determined by the flats have a line common to all. It is immediately obvious that, in the case in question: *The centres of the $n+2$ hyperspheres lie on a flat, viz. the common tangent flat to the hyperspheres at the point at infinity on their common line.*

In general, two real hyperspheres cannot have a single line in common. It follows that, if $n+2$ flats satisfy the condition in question, they cannot all be real; but the theorems may, of course, be projectively extended to the case of $n+2$ real quadrics which have a common section by a real flat.

In three dimensions, two spheres which have a common line touch at one finite point, generally unreal. It is easy to shew that, conversely: *If, of the circumspheres of the tetrahedra determined by five planes, any pair touch (at a point other than a vertex of one of the tetrahedra); then every pair touch, and the ten points of contact all lie on a line, which is a common generator of the five spheres.*

Digitized by Google

A mathematical theory of natural and artificial selection. Part III. By Mr J. B. S. HALDANE, Trinity College.

[*Received* 9 July, *read* 26 July 1926.]

In this part the cases of a single but incompletely dominant factor, and of several interacting factors are considered. Mating is supposed to be at random, populations to be very large, and generations not to overlap. The notation is, so far as possible, that of Part I (1).

Selection of an incompletely dominant autosomal character.

Let the nth generation be formed from female gametes in the ratio $u_n A : 1a$, male gametes in the ratio $v_n A : 1a$. The nth generation is therefore in the proportions $u_n v_n AA : (u_n + v_n) Aa : 1aa$. Let the ratios after selection has occurred be:

$$\begin{matrix} \text{♂} \\ \text{♀} \end{matrix} \quad \begin{matrix} u_n v_n AA : (1 - K_m)(u_n + v_n) Aa : (1 - k_m) aa, \\ u_n v_n AA : (1 - K_f)(u_n + v_n) Aa : (1 - k_f) aa, \end{matrix}$$

where K_m, K_f, k_m, k_f are small.

$$\therefore \quad u_{n+1} = \frac{2u_n v_n + (1 - K_f)(u_n + v_n)}{(1 - K_f)(u_n + v_n) + 2 - 2k_f}.$$

Hence, since $\dfrac{u_n - v_n}{u_n}$ is clearly small,

$$\Delta u_n = \frac{v_n - u_n}{2} + \frac{u_n (K_f u_n - K_f + k_f)}{1 + u_n}, \text{ approximately,}$$

and $$\Delta v_n = \frac{u_n - v_n}{2} + \frac{u_n (K_m u_n - K_m + k_m)}{1 + u_n}, \text{ approximately.}$$

Δu_n and Δv_n can be shewn to differ by a small quantity of the second order.

$$\therefore \quad \Delta u_n = \frac{u_n (K u_n - K + k)}{1 + u_n} \qquad \text{.....(1·0)}$$

where $$K = \tfrac{1}{2}(K_f + K_m); \; k = \tfrac{1}{2}(k_f + k_m).$$

Equilibrium can only occur when $\Delta u_n = 0$, i.e. u_n tends either to zero, infinity, or to $1 - \dfrac{k}{K}$. Hence for equilibrium to be possible $\dfrac{k}{K} < 1$. If K be positive, i.e. heterozygotes are at a disadvantage compared with pure dominants, then $\Delta u_n \gtrless 0$ according as $u_n \gtrless 1 - \dfrac{k}{K}$. Hence the equilibrium is unstable if it exists. If K

Digitized by Google

be negative the equilibrium is stable if it exists. We have thus three cases to consider. In each

$$\frac{du_n}{dn} = \frac{u_n(Ku_n - K + k)}{1 + u_n} \text{ approximately,}$$

and the proportion of recessives $y_n = (1 + u_n)^{-2}$.

(*a*) No equilibrium, $\frac{k}{K} > 1$.

$$\therefore \quad (k - K)n = \log_e u_n + \frac{k - 2K}{K} \log_e \left(\frac{Ku_n - K + k}{k}\right) \ldots (1{\cdot}1)$$

making the usual convention that $u_0 = 1$.

Hence the values of u_n lie between two geometrical series, and selection is therefore vastly more efficacious on recessives than when dominance is complete, as in equations 2·4 and 4·3 of Part I.

(*b*) Stable equilibrium, $k > K$, $0 > K$.

$$\therefore \quad (K - k)n = \log_e \left(\frac{u_n}{u_0}\right) + \frac{2K - k}{K} \log_e \left(\frac{Ku_n - K + k}{Ku_0 - K + k}\right) \ldots (1{\cdot}2).$$

We must take $u_0 \gtrless 1 - \frac{k}{K}$ according as $u_n \gtrless 1 - \frac{k}{K}$.

Here again successive values of u_n lie between two geometrical series, so that the population proceeds fairly rapidly towards equilibrium. As Fisher (2) has pointed out, such cases probably occur in nature in connexion with factors governing size, where the heterozygote is at an advantage as compared with either type of homozygote.

(*c*) Unstable equilibrium, $K > 0$, $K > k$.

The population proceeds towards homozygosis in one direction or the other. This case can hardly occur in nature, as any mutants, either in an AA or an aa population, would be weeded out while still few in number.

Selection of an incompletely dominant sex-linked character.

The female sex is throughout supposed to be homogametic; if the male is homogametic the argument is the same, *mutatis mutandis*. Let the nth generation be formed from ova in the ratio $u_nA : 1a$, female-producing spermatozoa in the ratio $v_nA : 1a$. Let the ratios of the nth generation after selection be:

$$\text{♀} \quad u_nv_nAA : (1 - K)(u_n + v_n)Aa : (1 - k)aa,$$
$$\text{♂} \quad u_nA : (1 - k')a,$$

where K, k and k' are small.

Digitized by Google

$$\therefore \quad u_{n+1} = \frac{2u_n v_n + (1-K)(u_n + v_n)}{(1-K)(u_n + v_n) + 2 - 2k},$$

$$v_{n+1} = \frac{u_n}{1-k'}.$$

$$\therefore \quad \Delta u_n = \Delta v_n = \frac{u_n}{3+3u_n}[(2K+k')u_n - 2K + 2k + k'],$$

approximately, ...(2·0)

and $u_n = v_n$, approximately.

Hence u_n tends to zero, infinity, or $\frac{2K-2k-k'}{2K+k'}$. Equilibrium is possible if $\frac{k+k'}{2K+k'} < \frac{1}{2}$. It is stable if $2K+k'$ be negative, unstable if this quantity be positive. In each case

$$\frac{du_n}{dn} = \frac{u_n(2Ku_n + k'u_n - 2K + 2k + k')}{3(1+u_n)}, \text{ approximately,}$$

and the proportion of recessive females is $(1+u_n)^{-2}$, of recessive males $(1+u_n)^{-1}$. Three cases occur.

(*a*) No equilibrium, $\frac{k+k'}{2K+k'} > \frac{1}{2}$.

$$\therefore \quad \frac{2k+k'-2K}{3} n = \log_e u_n + \frac{2k-4K}{2K+k'} \log_e \left(\frac{2Ku_n + k'u_n - 2K + 2k + k'}{2k + 2k'} \right)$$

.....(2·1),

putting $u_0 = 1$. Selection therefore proceeds much as according to equation 7·2 of Part I.

(*b*) Stable equilibrium, $0 > 2K + k'$, $2k + k' > 2K$.

$$\therefore \quad \frac{2K-2k-k'}{3} n = \log_e \left(\frac{u_n}{u_0} \right) + \frac{2k-4K}{2K+k'} \log_e \left(\frac{2Ku_n + k'u_n - 2K + 2k + k'}{2Ku_0 + k'u_0 - 2K + 2k + k'} \right)$$

.....(2·2)

where $u_0 \gtrless u_\infty$ according as $u_n \gtrless u_\infty$.

The results of Robertson (3) suggest that milk-yield in cattle depends on one or more sex-linked factors which act most effectively when heterozygous, besides autosomal factors. If so human effort in this case has given K a negative value, while k and k' are nearly zero. Hence an equilibrium should be reached.

(*c*) Unstable equilibrium, $2K + k' > 0$, $2K > 2k + k'$.

The population proceeds in one direction or the other to homozygosis. This case can hardly occur in nature.

Multiple factors.

Many cases exist in nature where several factors are needed to

Digitized by Google

ensure the appearance of a character. Thus in wheat Nilsson-Ehle (4) found that any one of three dominant factors will produce redness, that is to say a white plant must be a triple recessive. On the other hand Saunders (5) found that in Matthiola three dominant factors are needed for slight hoariness, four for complete hoariness, so that a hoary plant is a multiple dominant. In other cases the effects of factors are merely additive, and selection will act on each independently of the others. It will be shown later that linkage between factors, unless very strong, is unimportant. We shall therefore at first consider unlinked factors, and shall confine ourselves to the case of complete dominance.

Selection of a multiple autosomal recessive character.

Let $A_1, A_2, \ldots A_r, \ldots A_m$ be m autosomal dominant factors, each of which produces the same effect, so that the multiple recessive alone competes with the other genotypes. Let $1-k$ of this type survive for every one of the others. Let y_n be the proportion of multiple recessives in the nth generation, formed from gametes in the proportion ${}_ru_nA_r : 1a_r$, and similarly for the other factor pairs. Then the nth generation consists of zygotes in the ratios:

$${}_ru_n^2 A_r A_r : 2{}_ru_n A_r a_r : 1a_r a_r, \text{ etc.}$$

$$\therefore \quad y_n = \prod_{r=1}^{m} (1 + {}_ru_n)^{-2}.$$

Of the $a_r a_r$ zygotes only $y_n(1+{}_ru_n)^2$ are multiple recessives. Hence

$${}_ru_{n+1} = \frac{{}_ru_n(1+{}_ru_n)}{1+{}_ru_n - ky_n(1+{}_ru_n)^2}.$$

$$\therefore \quad \Delta_r u_n = ky_n\,{}_ru_n(1+{}_ru_n), \text{ approximately, if } k \text{ be small.}$$

Putting $x = kn$ we have approximately:

$$\left.\begin{aligned} \frac{d\,{}_1u_n}{dx} &= y_n\,{}_1u_n(1+{}_1u_n) \\ &\cdots\cdots\cdots\cdots \\ &\cdots\cdots\cdots\cdots \\ \frac{d\,{}_ru_n}{dx} &= y_n\,{}_ru_n(1+{}_ru_n) \\ &\cdots\cdots\cdots\cdots \\ &\cdots\cdots\cdots\cdots \end{aligned}\right\}.$$

To eliminate the u's from these m equations put $s_r = \frac{{}_ru_n}{1+{}_ru_n}$.

$$\therefore \quad \frac{ds_r}{dx} = y_n s_r. \qquad \therefore \quad \log s_r = \int y_n dx + \log a_r.$$

Digitized by Google

$\therefore$ $s_r = a_r s$, where a_r is an integration constant independent of n and given by the initial state of the population.

$$\left.\begin{aligned} \therefore \quad y_n &= \prod_{r=1}^{m} (1 - a_r s)^2 \\ kn = x &= \int \frac{ds}{s y_n} \end{aligned}\right\} \quad \text{.....(3·0).}$$

The latter equation is integrable, and the elimination of s gives the required relation between y_n and kn.

$$\frac{dy_n}{dx} = -2{y_n}^2 \sum_{r=0}^{m} {}_r u_n,$$

whereas if only one factor is concerned,

$$\frac{dy_n}{dx} = -2{y_n}^2 u_n.$$

Now comparing these rates for equal values of y_n in the two cases, we note that since ${y_n}^{-\frac{1}{2}} = 1 + u_n = \prod_{r=1}^{m} (1 + {}_r u_n)$ $\therefore$ $u_n > \sum_{r=1}^{m} {}_r u_n$. Hence selection is slower than in the case of a character determined by one factor only. When however dominants are very rare, or when one a_r greatly exceeds the rest, i.e. one recessive factor is far commoner than the others, selection proceeds at about the same rate in the two cases. It is slowest when all the a_r's are equal.

Selection of a multiple sex-linked recessive character.

If $A_1, A_2, \ldots A_r, \ldots A_m$ are sex-linked (the female being homogametic) the nth generation formed from eggs in the ratios ${}_r u_n A : 1a$, etc., and female-producing spermatozoa in the ratios ${}_r v_n A : 1a$, etc., while z_n is the proportion of multiple recessive males, y_n of such females, and k is the coefficient of selection.

$$\therefore \quad {}_r u_{n+1} = \frac{2\,{}_r u_n\,{}_r v_n + {}_r u_n + {}_r v_n}{{}_r u_n + {}_r v_n + 2 - 2k y_n (1 + {}_r u_n)}.$$

$$ {}_r v_{n+1} = \frac{{}_r u_n}{1 - k z_n (1 + {}_r u_n)}.$$

$$ y_n = \prod_{r=1}^{m} (1 + {}_r u_n)^{-1} (1 + {}_r v_n)^{-1}.$$

$$ z_n = \prod_{r=1}^{m} (1 + {}_r u_n)^{-1}.$$

$\therefore$ Approximately ${}_r u_n = {}_r v_n$, $y_n = {z_n}^2$.

$$\therefore \quad 3\Delta\, {}_r u_n = k\, {}_r u_n (1 + {}_r u_n)(2{z_n}^2 + z_n).$$

Digitized by Google

As above, putting $kn = x$, $\frac{{}_ru_n}{1 + {}_ru_n} = a_r s$, we have a_r constant, and

$$\left.\begin{aligned} z_n &= \prod_{r=1}^{m} (1 - a_r s) \\ kn = x &= 3 \int \frac{ds}{s z_n (2z_n + 1)} \end{aligned}\right\} \quad \text{.....(4·0).}$$

This again is soluble in finite terms by the elimination of s.

$$\frac{dz_n}{dx} = -\tfrac{1}{3} z_n^2 (2z_n + 1) \sum_{r=1}^{m} {}_ru_n,$$

while in the single factor case

$$\frac{dz_n}{dx} = -\tfrac{1}{3} z_n^2 (2z_n + 1)\, u_n, \text{ where } u_n = z_n^{-1} - 1.$$

Comparing these rates for equal values of z_n, we find as above $u_n > \Sigma\, {}_ru_n$. Hence selection proceeds more slowly with many factors than with one. When, however, dominants are very rare or one a_r much larger than the rest, selection proceeds as with one factor.

Selection of a multiple autosomal dominant character.

When each of m autosomal dominant factors is needed to produce a character, we find, using the same notation as above except that y_n is the proportion of dominants,

$$y_n = \prod_{r=1}^{m} [1 - (1 + {}_ru_n)^{-2}],$$

$$\frac{d\, {}_ru_n}{dx} = \frac{y_n (1 + {}_ru_n)}{2 + {}_ru_n}, \text{ with } m - 1 \text{ similar equations.}$$

$$\therefore \quad (1 + {}_ru_n)\, e^{1 + {}_ru_n} = e^{\int y_n dx} = a_r s.$$

Hence the problem can be reduced to the elimination of s between:

$$\left.\begin{aligned} y_n &= \prod_{r=1}^{m} [1 - \{\phi (a_r s)\}^{-2}] \\ kn = x &= \int \frac{ds}{sy} \end{aligned}\right\} \quad \text{.....(5·0)}$$

where ϕ is defined by the equation $t = \phi(t)\, e^{\phi(t)}$.

Numerical integration would be possible for known values of a_r,

$$\frac{dy_n}{dx} = 2y_n^2 \sum_{r=1}^{m} \frac{1}{{}_ru_n (2 + {}_ru_n)^2},$$

while in the single factor case

$$\frac{dy_n}{dx} = \frac{2y_n^2}{u_n (2 + u_n)^2}.$$

Now when one a_r is very much smaller than the rest these two

Digitized by Google

rates are nearly equal for equal values of y_n. When all the a_r's are equal,

$$\frac{dy_n}{dx} = my_n^2\left[\left(1-y_n^{\frac{1}{m}}\right)^{-\frac{1}{2}}+1\right]^{-2}\left[\left(1-y_n^{\frac{1}{m}}\right)^{-2}-1\right]^{-1}.$$

The ratio of this rate to the rate with a single factor (putting $t^m = y_n$) is

$$\frac{mt^{m-1}(1+\sqrt{1-t^m})(1-t)^{\frac{3}{2}}}{(1+\sqrt{1-t})(1-t^m)^{\frac{3}{2}}}.$$

When t is small this tends to the small mt^{m-1}; when t is nearly unity, to $m^{-\frac{1}{2}}$ which is also small. The ratio when all the a_r's are equal is, by Purkiss' theorem, the minimum value. Hence it would seem that in general natural selection acts more slowly on a multiple dominant than a single dominant. The case of a multiple sex-linked dominant and various more complicated cases present still greater difficulties to analysis, though of course individual cases could always be solved numerically.

Linkage.

Consider two autosomal factors A, B, linked with such intensity that the cross-over value is $100l$ in the female, $100l'$ in the male sex. Let the nth generation be formed from:—

eggs $\quad p_nAB : q_nAb : r_naB : s_nab,$

spermatozoa $\quad p_n'AB : q_n'Ab : r_n'aB : s_n'ab,$

where $p_n + q_n + r_n + s_n = p_n' + q_n' + r_n' + s_n' = 1.$

The nth generation therefore consists of:—

$$p_np_n'ABAB : (p_nq_n' + p_n'q_n)ABAb : (p_nr_n' + p_n'r_n)ABaB$$
$$: q_nq_n'AbAb : (p_ns_n' + p_n's_n)AB.ab : (q_nr_n' + q_n'r_n)Ab.aB$$
$$: r_nr_n'aB.aB : (q_ns_n' + q_n's_n)Abab : (r_ns_n' + r_n's_n)aBab$$
$$: s_ns_n'abab.$$

If no selection occurs they produce gametes in the proportions:

$$2p_{n+1} = p_n + p_n' + l(q_nr_n' + q_n'r_n - p_ns_n' - p_n's_n)$$
$$2q_{n+1} = q_n + q_n' - l(q_nr_n' + q_n'r_n - p_ns_n' - p_n's_n)$$
$$2r_{n+1} = r_n + r_n' - l(q_nr_n' + q_n'r_n - p_ns_n' - p_n's_n)$$
$$2s_{n+1} = s_n + s_n' + l(q_nr_n' + q_n'r_n - p_ns_n' - p_n's_n),$$

whilst the values of p_{n+1}', etc., are given by the same expressions with l' substituted for l. Hence after one generation

$$\frac{p_n+q_n}{r_n+s_n} \text{ and } \frac{p_n'+q_n'}{r_n'+s_n'}$$

have the same constant value u, while

$$\frac{p_n+r_n}{q_n+s_n} = \frac{p_n'+r_n'}{q_n'+s_n'} = v.$$

Digitized by Google

We may therefore write:

$$p_n = \frac{uv}{(1+u)(1+v)} + x_n; \quad q_n = \frac{u}{(1+u)(1+v)} - x_n;$$

$$r_n = \frac{v}{(1+u)(1+v)} - x_n; \quad s_n = \frac{1}{(1+u)(1+v)} + x_n.$$

$$p_n' = \frac{uv}{(1+u)(1+v)} + x_n'; \quad q_n' = \frac{u}{(1+u)(1+v)} - x_n';$$

$$r_n' = \frac{v}{(1+u)(1+v)} - x_n'; \quad s_n' = \frac{1}{(1+u)(1+v)} + x_n'.$$

$$\therefore \quad q_n r_n' + q_n' r_n - p_n s_n' - p_n' s_n = -2(x_n + x_n').$$

$$\therefore \quad 2x_{n+1} = (1-2l)(x_n + x_n'); \; 2x_{n+1}' = (1-2l')(x_n + x_n').$$

Hence if $x_0 + x_0' = c$,

$$\left.\begin{aligned} \therefore \quad x_n &= (\tfrac{1}{2} - l)(1-l-l')^{n-1}c \\ x_n' &= (\tfrac{1}{2} - l')(1-l-l')^{n-1}c \end{aligned}\right\} \quad \text{.....(6·0).}$$

Hence the proportions of the various types of gamete approach asymptotically those which would be reached in one generation without linkage, the ratio of successive differences from the final values being $1-l-l'$. Hence if either l or l' is larger than k the effects of linkage are unimportant. A similar proof holds for a pair of sex-linked factors.

Selection in a tetraploid organism.

In a tetraploid race which is stable, i.e. yields only diploid gametes, five types of zygote and three of gamete exist. Gregory (6) and Blakeslee, Belling and Farnham (7) have shown that zygotes produce gametes as follows:

Zygotes	Gametes
AAAA	*AA*
AAAa	1*AA* : 1*Aa*
AAaa	1*AA* : 4*Aa* : 1*aa*
Aaaa	1*Aa* : 1*aa*
aaaa	*aa*

Gregory thought that *AAaa* gave 1*AA* : 2*Aa* : 1*aa*, but his results, as well as theory, suggest the above ratio. As in Part II we first consider tetraploidy without selection, and then the process of selection in a population which would be in equilibrium but for that selection. Let the mth generation be formed from gametes in the ratios $p_m AA : 2q_m Aa : r_m aa$, where $p_m + 2q_m + r_m = 1$, and $u_m = \dfrac{p_m + q_m}{q_m + r_m}$. They form zygotes in the ratios:

$$p_m^2 AAAA : 4p_m q_m AAAa : (4q_m^2 + 2p_m r_m) AAaa$$
$$: 4q_m r_m Aaaa : r_m^2 aaaa.$$

Digitized by Google

$$\therefore \quad p_{m+1} = p_m + \tfrac{2}{3}(q_m^2 - p_m r_m)$$
$$q_{m+1} = q_m - \tfrac{2}{3}(q_m^2 - p_m r_m)$$
$$r_{m+1} = r_m + \tfrac{2}{3}(q_m^2 - p_m r_m).$$

Hence $u_{m+1} = u_m = u$, and when equilibrium is reached $q_\infty^2 = p_\infty r_\infty$.

Hence $p_\infty = \dfrac{u^2}{(1+u)^2}$, $q_\infty = \dfrac{u}{(1+u)^2}$, $r_\infty = \dfrac{1}{(1+u)^2}$,

and the population in equilibrium is in the ratios:

$$\frac{u^4}{(1+u)^4} AAAA : \frac{4u^3}{(1+u)^4} AAAa : \frac{6u^2}{(1+u)^4} AAaa$$
$$: \frac{4u}{(1+u)^4} Aaaa : \frac{1}{(1+u)^4} aaaa.$$

Putting $\theta_m = q_m^2 - p_m r_m$, we find $\theta_{m+1} = \frac{1}{3}\theta_m$, $\therefore \theta_m = 3^{-m}\theta_0$.

$$\therefore \quad p_m = p_{m-1} + \theta_{m-1}$$
$$= p_0 + \sum_{r=0}^{m-1} \theta_r$$
$$= p_0 + \tfrac{3}{2}(1 - 3^{-m})\theta_0$$
$$= p_\infty - \frac{3^{1-m}}{2}\theta_0 \qquad \text{.....(7·0).}$$

Hence the ratios of the different classes converge very rapidly to their final values. Under selection of a population which has reached such an equilibrium, if A is completely dominant,

$$u_{n+1} = \frac{p_n + q_n}{q_n + r_n - k r_n^2}$$
$$= \frac{u_n}{1 - k(1+u_n)^{-3}},$$

$$\therefore \quad \frac{du_n}{dn} = \Delta u_n = \frac{k u_n}{(1+u_n)^3}, \text{ approximately, if } k \text{ be small.}$$

$$\therefore \text{ if } u_0 = 1, \quad kn = \log_e u_n + 3u_n + \tfrac{3}{2}u_n^2 + \tfrac{1}{3}u_n^3 - 4\tfrac{5}{6} \ldots (7{\cdot}1),$$

Hence when dominants are few u_n changes at the same rate as in a diploid organism; when they are many, much more slowly. To compare the change in the number y_n of recessives we find

$$kn = \log_e(1 - y_n^{\frac{1}{4}}) - \tfrac{1}{4}\log_e y_n + y_n^{-\frac{1}{4}} + \tfrac{1}{2}y_n^{-\frac{1}{2}} + \tfrac{1}{3}y_n^{-\frac{3}{4}} - 6\tfrac{2}{3} \ldots (7{\cdot}2),$$

$$\therefore \quad \frac{dy_n}{dn} = -4k y_n^2 (y_n^{-\frac{1}{4}} - 1),$$

while in a diploid population

$$\frac{dy_n}{dn} = -2k y_n^2 (y_n^{-\frac{1}{2}} - 1).$$

Hence here too the rate is always slower in the tetraploids, though not much so when recessives are few.

Digitized by Google

If dominance is incomplete, as is usual in tetraploid organisms, and after selection the zygotes are in the ratios:

$$u_n^4 AAAA : 4(1-k_1)u_n^3 AAAa : 6(1-k_2)u_n^2 AAaa$$
$$: 4(1-k_3)u_n Aaaa : (1-k_4)aaaa,$$

$$\therefore \quad \Delta u_n = \frac{k_1 u_n^4 + 3(k_2-k_1)u_n^3 + 3(k_3-k_2)u_n^2 + (k_4-k_3)u_n}{(1+u_n)^3},$$

approximately, if the coefficients are small. The possible equilibria, if any, are given by the roots of

$$k_1 u_\infty^3 + 3(k_2-k_1)u_\infty^2 + 3(k_3-k_2)u_\infty + k_4 - k_3 = 0.$$

The various possible cases, and their stability, could easily be investigated. If the advantage of the various genotypes increases or decreases with the number of dominant factors they contain, so that $k_4 > k_3 > k_2 > k_1 > 0$, or $0 > k_1 > k_2 > k_3 > k_4$, no equilibrium is possible,

$$\therefore \quad n = \int \frac{(1+u_n)^3 du_n}{k_1 u_n^4 + 3(k_2-k_1)u_n^3 + 3(k_3-k_2)u_n^2 + (k_4-k_3)u_n} \quad \ldots(7{\cdot}3).$$

If $k_1 = 0$ this contains a term proportional to u_n or u_n^2. If $k_1 \neq 0$ all the terms are logarithmic and selection is always rapid. But $AAAa$ is more likely to resemble $AAAA$ than Aa to resemble AA. Hence polyploidy diminishes the probability of a rapid selection in populations where recessives are few. Since stable polyploidy is only known in hermaphrodite plants there is no need to discuss cases of sex-linkage or different intensities of selection in the two sexes. The theory can readily be extended to the higher forms of polyploidy.

Summary.

Expressions are found for the changes caused by slow selection in populations whose characters are determined by incompletely dominant, multiple, or polyploid factors, and for the equilibria attained in certain of these cases.

REFERENCES.

1. J. B. S. Haldane. *Trans. Camb. Phil. Soc.*, vol. 22, p. 19, 1924.
2. R. A. Fisher. *Proc. Roy. Soc. Edin.*, vol. 42, p. 321, 1922.
3. E. Robertson. *Journ. Genetics*, vol. 11, p. 79, 1921.
4. H. Nilsson-Ehle. *Lund's Univ. Arsskrift.*, p. 69, 1909.
5. E. R. Saunders. *Journ. Genetics*, vol. 10, p. 149, 1920.
6. Gregory. *Proc. Roy. Soc.*, B, vol. 87, p. 484, 1914.
7. Blakeslee, Belling and Farnham. *Science*, vol. 52, p. 388, 1920.

Digitized by Google

The Summation of a Fourier Integral of Finite Type. By Mr S. POLLARD, Trinity College.

[*Received* 14 August, *read* 25 October 1926.]

1. A Fourier integral will be described as of *finite type* if the range of integration of the integrals by means of which the coefficients in the Fourier integral are defined is a finite interval instead of the usual $(-\infty, \infty)$. Thus

$$\int_0^\infty (a_s \cos sx + b_s \sin sx)\, ds \quad \ldots\ldots\ldots\ldots\ldots\ldots (1{\cdot}1)$$

is of finite type if there is a finite interval (p, q), such that

$$a_s = \frac{1}{\pi}\int_p^q f(t)\cos st\,dt, \quad b_s = \frac{1}{\pi}\int_p^q f(t)\sin st\,dt, \ldots (1{\cdot}2),$$

where $f(x)$ is the generating function of the series.

Any Fourier integral is of finite type if the generating function vanishes for all sufficiently large values of x. Thus, if $f(x) = 0$ outside (p, q), the usual formulae

$$a_s = \frac{1}{\pi}\int_{-\infty}^\infty f(t)\cos st\,dt, \quad b_s = \frac{1}{\pi}\int_{-\infty}^\infty f(t)\sin st\,dt,$$

are equivalent to the above.

It is easily seen that any Fourier integral can be expressed as the sum of two, one of which is of finite type and the other of which is of null type, i.e. is such that its generating function vanishes in some neighbourhood of the point at which the Fourier integral is considered. When the generating function is not absolutely integrable in $(-\infty, \infty)$ this division of the Fourier integral has much to recommend it as the difficulties presented by integrals of the one type are unlike those of the other type, and it only causes confusion to attempt to discuss them together.

In the present paper it will be supposed that $f(x)$ is integrable in the general Denjoy sense in (p, q) where p and q are finite, and that the coefficients a_s, b_s are defined by (1·2). To make matters quite definite it will also be supposed that $f(x) = 0$ outside (p, q).

The problem considered is that of the summation at a point x in $p < x < q$ of the integral (1·1), i.e. the behaviour of the integrals

$$I_k(\omega) = \int_0^\omega \left(1 - \frac{s}{\omega}\right)^k (a_s \cos sx + b_s \sin sx)\, ds$$

as $\omega \to \infty$.

Digitized by Google

It will be supposed throughout that $k \geqslant 1$. Under this condition it will be shown that the Fourier integral behaves in exactly the same way as the corresponding series.

2. A simple calculation shows that

$$a_s \cos sx + b_s \sin sx = \frac{1}{\pi} \int_p^q f(t) \cos s(t-x)\, dt,$$

so that $$I_k(\omega) = \frac{1}{\pi} \int_0^\omega \left(1 - \frac{s}{\omega}\right)^k ds \int_p^q f(t) \cos s(t-x)\, dt.$$

Before further progress can be made the order of integration in this integral must be inverted. Choose, as is plainly possible, a constant γ such that, when $g(t) = f(t) - \gamma$,

$$\int_p^q g(t)\, dt = 0,$$

and write $$G(t) = \int_p^t g(u)\, du.$$

Then

$$\int_p^q g(t) \cos s(t-x)\, dt = \Big[G(t) \cos s(t-x)\Big]_p^q + s \int_p^q G(t) \sin s(t-x)\, dt$$

$$= s \int_p^q G(t) \sin s(t-x)\, dt,$$

since $G(q) = G(p) = 0$, and

$$I_k(\omega) = \frac{\gamma}{\pi} \int_0^\omega \left(1 - \frac{s}{\omega}\right)^k ds \int_p^q \cos s(t-x)\, dt$$

$$+ \frac{1}{\pi} \int_0^\omega s \left(1 - \frac{s}{\omega}\right)^k ds \int_p^q G(t) \sin s(t-x)\, dt.$$

Now $G(t)$, as an indefinite integral, is a continuous function. Thus in both the integrals on the right the integrand is continuous, and the order of integration may be inverted, so that

$$I_k(\omega) = \frac{1}{\pi} \int_p^q \gamma\, dt \int_0^\omega \left(1 - \frac{s}{\omega}\right)^k \cos s(t-x)\, ds$$

$$+ \frac{1}{\pi} \int_p^q G(t)\, dt \int_0^\omega s \left(1 - \frac{s}{\omega}\right)^k \sin s(t-x)\, ds.$$

Since

$$\int_0^\omega s \left(1 - \frac{s}{\omega}\right) \sin s(t-x)\, ds = -\frac{d}{dt} \int_0^\omega \left(1 - \frac{s}{\omega}\right) \cos s(t-x)\, ds,$$

Digitized by Google

integration by parts shows that

$$\int_p^q G(t)\,dt \int_0^\omega s\left(1-\frac{s}{\omega}\right)^k \sin s(t-x)\,ds$$

$$= \left[-G(t)\int_0^\omega \left(1-\frac{s}{\omega}\right)^k \cos s(t-x)\,ds\right]_p^q$$

$$+ \int_p^q g(t)\,dt \int_0^\omega \left(1-\frac{s}{\omega}\right)^k \cos s(t-x)\,ds$$

$$= \int_p^q g(t)\,dt \int_0^\omega \left(1-\frac{s}{\omega}\right)^k \cos s(t-x)\,ds.$$

Thus

$$I_k(\omega) = \frac{1}{\pi}\int_p^q f(t)\,dt \int_0^\omega \left(1-\frac{s}{\omega}\right)^k \cos s(t-x)\,ds$$

$$= \frac{1}{\pi}\int_{p-x}^{q-x} f(x+t)\,dt \int_0^\omega \left(1-\frac{s}{\omega}\right)^k \cos ts\,ds.$$

3. The function

$$\int_0^\omega \left(1-\frac{s}{\omega}\right)^k \cos ts\,ds$$

can be expressed in terms of functions which have been investigated by various writers in some detail. It may be rewritten in the form

$$\omega\gamma_{k+1}(\omega t)$$

where

$$\gamma_{k+1}(t) = \int_0^1 (1-u)^k \cos tu\,du$$

$$= \frac{\Pi(k)}{t^{k+1}} C_{k+1}(t) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3{\cdot}1),$$

the function $C_{k+1}(t)$ being the case $p = k+1$ of

$$C_p(t) = \sum_{r=0}^{\infty} (-)^r \frac{t^{p+2r}}{\Pi(p+2r)}.$$

This last-mentioned function has the convenient property

$$\frac{d}{dt} C_{p+1}(t) = C_p(t) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3{\cdot}2);$$

also for $0 \leqslant p \leqslant 2$ it is bounded, while for $p > 2$,

$$C_p(t) \sim \frac{t^{p-2}}{\Pi(p-2)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3{\cdot}3).$$

Digitized by Google

From this last, it follows that

$$|\gamma_{k+1}(t)| \leqslant \frac{A}{t^{k+1}} \quad (0 \leqslant k \leqslant 1) \quad \ldots\ldots(3{\cdot}4),$$

$$\gamma_{k+1}(t) \sim \frac{k}{t^2} \quad (k \geqslant 1) \quad \ldots\ldots\ldots\ldots(3{\cdot}5).$$

If r is any positive integer not greater than $k+1$, (3·2) together with Leibniz' theorem yields

$$\frac{1}{\Pi(k)}\gamma_{k+1}^{(r)}(t) = \frac{C_{k-r+1}(t)}{t^{k+1}} - r(k+1)\frac{C_{k-r+2}(t)}{t^{k+2}} + \ldots + (-)^r(k+1)\ldots(k+r)\frac{C_{k+1}(t)}{t^{k+r+1}}.$$

Thus, by (3·3) and what precedes, as $t \to \infty$,

$$\gamma_{k+1}^{(r)}(t) = O\left(\frac{1}{t^{r+2}}\right) \quad (r \leqslant k-1) \quad \ldots(3{\cdot}6),$$

$$= O\left(\frac{1}{t^{k+1}}\right) \quad (r \geqslant k-1) \quad \ldots(3{\cdot}7).$$

In the special case $r = k$, as $t \to \infty$,

$$\frac{1}{\Pi(k)}\gamma_{k+1}^{(k)}(t) = \frac{C_1(t)}{t^{k+1}} + O\left(\frac{1}{t^{k+2}}\right) = \frac{\sin t}{t^{k+1}} + O\left(\frac{1}{t^{k+2}}\right).$$

For small values of t it is convenient to proceed from the formula

$$\gamma_{k+1}^{(r)}(t) = \int_0^1 (1-u)^k u^r \cos^{(r)} tu\, du \ldots\ldots\ldots\ldots(3{\cdot}8),$$

which yields

$$|\gamma_{k+1}^{(r)}(t)| \leqslant \int_0^1 (1-u)^k u^r\, du = B(k+1, r+1).$$

It is convenient to write this in the form

$$|\gamma_{k+1}^{(r)}(t)| \leqslant A \quad \ldots\ldots\ldots\ldots\ldots\ldots(3{\cdot}9),$$

where, with the usual notation, A is constant relative to each function considered but may vary from function to function.

4. For convenience the suffix $k+1$ will now be omitted, and $\gamma(t)$ simply written instead of $\gamma_{k+1}(t)$. This makes the formulae easier to read and causes no confusion as k remains unchanged throughout each set of calculations.

Digitized by Google

If $\bar{\eta}$ is any positive number less than both $q-x$ and $x-p$, then

$$\int_{\eta}^{q-x} f(x+t).\,\omega\gamma(\omega t)\,dt$$

$$= \Big[F(x+t).\,\omega\gamma(\omega t)\Big]_{\eta}^{q-x} - \int_{\eta}^{q-x} F(x+t).\,\omega^2\gamma'(\omega t)\,dt,$$

where $F(x)$ is an indefinite integral of $f(x)$.

As the integrals employed are Denjoy integrals, the use of the formula for integration by parts requires justification. Here it is valid because $\gamma'(\omega t)$ is bounded, and therefore $\gamma(\omega t)$ is of bounded variation.

In the range of integration

$$|\omega\gamma(\omega t)| \leqslant \frac{A}{\omega^k \eta^{k+1}},$$

so that the term between the limits $\to 0$ with $\frac{1}{\omega^k}$, since $k \geqslant 1$.

For the remaining term it may be observed that, if (λ, μ) is any interval in the range $(\eta, q-x)$, then

$$\int_{\lambda}^{\mu} \omega^2\gamma'(\omega t)\,dt = \Big[\omega\gamma(\omega t)\Big]_{\lambda}^{\mu} \to 0.$$

It follows, by Lebesgue's convergence theorem, since $F(x+t)$ is continuous, that

$$\int_{\eta}^{q-x} F(x+t).\,\omega^2\gamma'(\omega t)\,dt \to 0$$

as $\omega \to \infty$.

Thus $$\int_{\eta}^{q-x} f(x+t).\,\omega\gamma(\omega t)\,dt \to 0 \qquad \text{.....(4·1).}$$

Similarly $$\int_{p-x}^{-\eta} f(x+t).\,\omega\gamma(\omega t)\,dt \to 0 \qquad \text{.....(4·2)}$$

and, as $\omega \to \infty$, the limits of $I_k(\omega)$ are identical with those of

$$\frac{1}{\pi}\int_{-\eta}^{\eta} f(x+t).\,\omega\gamma(\omega t)\,dt.$$

It is convenient to write this last expression in the form

$$\frac{\omega}{\pi}\int_{0}^{\eta} \psi(t)\,\gamma(\omega t)\,dt,$$

where $$\psi(t) = f(x+t) + f(x-t).$$

5. It is easily verified that, as $\omega \to \infty$,

$$\frac{\omega}{\pi}\int_{0}^{\eta} \gamma(\omega t)\,dt \to \tfrac{1}{2},$$

Digitized by Google

so that the necessary and sufficient condition for summability (C, k) to s at x is

$$\omega \int_0^\eta \phi(t)\,\gamma(\omega t)\,dt \to 0,$$

where $$\phi(t) = f(x+t) + f(x-t) - 2s.$$

In view of (4·1) and (4·2) the limits of this last-mentioned integral are independent of η, so that

$$\operatorname*{limm}_{\omega\to\infty} \omega \int_0^\eta \phi(t)\,\gamma(\omega t)\,dt = \lim_{\eta\to 0} \operatorname*{limm}_{\omega\to\infty} \omega \int_0^\eta \phi(t)\,\gamma(\omega t)\,dt,$$

on the principle that a constant is its own limit. Thus the necessary and sufficient condition takes the form

$$\lim_{\eta\to 0} \overline{\lim_{\omega\to\infty}} \left| \omega \int_0^\eta \phi(t)\,\gamma(\omega t)\,dt \right| = 0 \quad \text{.....(5·1),}$$

which is identical with the corresponding condition for Fourier series.

6. Write

$$\Phi_1(t) = \int_0^t \phi(u)\,du,$$

$$\Phi_2(t) = \int_0^t \Phi_1(u)\,du,$$

.....

Repeated integration by parts gives

$$\begin{aligned}
&\omega \int_0^\eta \phi(t)\,\gamma(\omega t)\,dt \\
&= \left[\omega\Phi_1(t)\,\gamma(\omega t) - \omega^2\Phi_2(t)\,\gamma'(\omega t) + \ldots + (-1)^r \omega^r \Phi_r(t)\,\gamma^{(r-1)}(\omega t) \right]_0^\eta \\
&\qquad + (-1)^{r+1}\omega^{r+1} \int_0^\eta \Phi_r(t)\,\gamma^{(r)}(\omega t)\,dt \\
&= \omega\Phi_1(\eta)\,\gamma(\omega\eta) - \omega^2\Phi_2(\eta)\,\gamma'(\omega\eta) + \ldots + (-1)^r \omega^r \Phi_r(\eta)\,\gamma^{(r-1)}(\omega\eta) \\
&\qquad + (-1)^{r+1}\omega^{r+1} \int_0^\eta \Phi_r(t)\,\gamma^{(r)}(\omega t)\,dt.
\end{aligned}$$

If $r \leqslant k$, then (3·6) and (3·7) show that

$$\omega^{p+1}\gamma^{(p)}(\omega\eta) = O\left(\frac{1}{\omega}\right) \quad (p = 1, 2, \ldots, r-1).$$

Also, by (3·5), $$\omega\gamma(\omega\eta) = O\left(\frac{1}{\omega}\right).$$

Digitized by Google

Thus

$$\omega \int_0^\eta \phi(t)\,\gamma(\omega t)\,dt = O\left(\frac{1}{\omega}\right) + (-1)^{r+1}\omega^{r+1}\int_0^\eta \Phi_r(t)\,\gamma^{(r)}(\omega t)\,dt,$$

and (5·1) is equivalent to

$$\lim_{\eta\to 0}\overline{\lim_{\omega\to\infty}}\left|\omega^{r+1}\int_0^\eta \Phi_r(t)\,\gamma^{(r)}(\omega t)\,dt\right| = 0 \qquad \text{.....(6·1),}$$

which is accordingly another form of the necessary and sufficient condition for summability (C, k) to s at x.

7. A sufficient condition is

$$\phi_r^*(t)\to 0 \qquad \text{.....(7·1),}$$

as $t\to 0$, where $\phi_r^*(t) = \dfrac{1}{t}\displaystyle\int_0^t \frac{|\Phi_{r-1}(u)|}{u^{r-1}}\,du$

and r is any positive integer less than $k+1$.

For $r-1\leqslant k$ and so

$$\lim_{\eta\to 0}\overline{\lim_{\omega\to\infty}}\left|\omega^r\int_0^\eta \Phi_{r-1}(t)\,\gamma^{(r-1)}(\omega t)\,dt\right| = 0$$

is a valid form of (6·1).

Now

$$\left|\omega^r\int_0^\eta \Phi_{r-1}(t)\,\gamma^{(r-1)}(\omega t)\,dt\right|$$

$$\leqslant \left|\omega^r\int_0^{1/\omega}\Phi_{r-1}(t)\,\gamma^{(r-1)}(\omega t)\,dt\right| + \left|\omega^r\int_{1/\omega}^\eta \Phi_{r-1}(t)\,\gamma^{(r-1)}(\omega t)\,dt\right|$$

$$\leqslant A\omega^r\int_0^{1/\omega}|\Phi_{r-1}(t)|\,dt + \omega^r\int_{1/\omega}^\eta |\Phi_{r-1}(t)|\,\frac{A'}{(\omega t)^{r+\epsilon}},$$

by (3·9) and (3·6), where $\epsilon = min\,(1,\ k+1-r)$,

$$\leqslant \frac{A}{\omega}\int_0^{1/\omega}\frac{|\Phi_{r-1}(t)|}{t^{r-1}}\,dt + \frac{A'}{\omega^\epsilon}\int_{1/\omega}^\eta \frac{|\Phi_{r-1}(t)|}{t^{r-1}}\,\frac{dt}{t^{1+\epsilon}}.$$

In this last expression the first term is none other than

$$A\phi_r^*\left(\frac{1}{\omega}\right),$$

which tends to zero by (7·1), while the second becomes on integration by parts

$$\frac{A'}{\omega^\epsilon}\left[\frac{\phi_r^*(t)}{t^\epsilon}\right]_{1/\omega}^\eta + \frac{(1+\epsilon)A'}{\omega^\epsilon}\int_{1/\omega}^\eta \phi_k^*(t)\,\frac{dt}{t^{1+\epsilon}}$$

$$\leqslant \frac{A'\phi_r^*(\eta)}{(\omega\eta)^\epsilon} + A'\phi_r^*\left(\frac{1}{\omega}\right) + \frac{(1+\epsilon)A'M(\eta)}{\omega^\epsilon}\int_{1/\omega}^\eta \frac{dt}{t^{1+\epsilon}}$$

Digitized by Google

where $M(\eta)$ is the upper bound of $\phi_r^*(t)$ in $(0, \eta)$, and tends to zero by (7·1).

The first two terms plainly tend to zero as $\omega \to \infty$. The last term may be written

$$\frac{(1+\epsilon)A'M(\eta)}{\epsilon}\left\{1 - \frac{1}{(\omega\eta)^{\epsilon}}\right\},$$

and is annihilated by the operator $\lim_{\eta \to 0} \overline{\lim_{\omega \to \infty}}$.

Thus (7·1) is sufficient for summability (C, k) whenever $r < k+1$. In particular, when k is integral,

$$\phi_k^*(t) \to 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (7{\cdot}2)$$

is sufficient.

8. A special case of (7·2) is

$$\frac{\Phi_{k-1}(t)}{t^{k-1}} \to 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (8{\cdot}1),$$

when (7·2) follows without difficulty.

Now (8·1) is also the condition that s is the limit $(C, k-1)$ of $f(u)$ as $u \to x$, so that a sufficient condition for summability (C, k), where k is integral, to s is

$$f(u) \to s \qquad (C, k-1).$$

The analysis in §§ 6, 7 may thus be interpreted as a direct proof, for Fourier integrals of the type considered, and also for Denjoy-Fourier series, of Hardy's sufficient condition for summability (C). Thus it is sufficient for summability (C), i.e. summability by some Cesàro mean, that $f(u)$ have a finite limit (C) as $u \to x$.

9. Suppose still that k is integral and that

$$\phi_k(t) \to 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (9{\cdot}1),$$

where $$\phi_k(t) = \frac{\Phi_k(t)}{t^k}.$$

This new condition is much less restricted than (7·2), and it is highly probable that by itself it is not sufficient for summability (C, k).

The necessary and sufficient condition for summability (C, k) can be taken in the form

$$\lim_{\eta \to 0} \overline{\lim_{\omega \to \infty}} \left| \omega^{k+1} \int_0^{\eta} \Phi_k(t)\, \gamma^{(k)}(\omega t)\, dt \right| = 0.$$

Digitized by Google

Now

$$\left|\omega^{k+1}\int_0^{1/\omega}\Phi_k(t)\,\gamma^{(k)}(\omega t)\,dt\right| \leqslant A\omega\int_0^{1/\omega}\frac{|\Phi_k(t)|}{t^k}\,dt$$

$$= A\omega\int_0^{1/\omega}|\phi_k(t)|\,dt,$$

which is known to tend to zero if $\Phi_k(t)\to 0$. Thus, when (9·1) holds the necessary and sufficient condition for summability (C, k) is

$$\lim_{\eta\to 0}\overline{\lim_{\omega\to\infty}}\left|\omega^{k+1}\int_{1/\omega}^{\eta}\Phi_k(t)\,\gamma^{(k)}(\omega t)\,dt\right| = 0 \quad \ldots\ldots(9\cdot 2).$$

By § 3, $\gamma^{(k)}(\omega t)$ can be written in the form

$$A\frac{\sin t}{t^{k+1}} + O\left(\frac{1}{t^{k+2}}\right).$$

Now $$\omega^{k+1}\int_{1/\omega}^{\eta}\Phi_k(t)\frac{dt}{(\omega t)^{k+2}} = \frac{1}{\omega}\int_{1/\omega}^{\eta}\phi_k(t)\,\frac{dt}{t^2},$$

which is not greater in absolute value than

$$\frac{\mu(\eta)}{\omega}\int_{1/\omega}^{\eta}\frac{dt}{t^2} = \mu(\eta) - \frac{\mu(\eta)}{\eta\omega},$$

where $\mu(\eta)$ is the upper bound of $\phi_k(t)$ in $(0, \eta)$.

This last expression is annihilated by the operator $\lim_{\eta\to 0}\overline{\lim_{\omega\to\infty}}$ and it follows that (9·2) is equivalent to

$$\lim_{\eta\to 0}\overline{\lim_{\omega\to\infty}}\left|\omega^{k+1}\int_{1/\omega}^{\eta}\Phi_k(t)\,\frac{\sin\omega t}{(\omega t)^{k+1}}\,dt\right| = 0,$$

which may be rewritten as

$$\lim_{\eta\to 0}\overline{\lim_{\omega\to\infty}}\left|\int_{1/\omega}^{\eta}\phi_k(t)\,\frac{\sin\omega t}{t}\,dt\right| = 0.$$

This, however, since $\phi_k(t)\to 0$, is none other than the necessary and sufficient condition that the Fourier series of $\phi_k(t)$ be convergent to 0 at $t=0$. Thus, when $\phi_k(t)\to 0$ the problem of summability (C, k) is reduced to that of convergence in the ordinary way.

10. A sufficient condition for summability (C, k) to s is

$$\phi_k(t) = o(1), \quad \phi_k(t) = O(1) \quad \ldots\ldots(10\cdot 1).$$

For, by a test due to W. H. Young, the Fourier series of $\phi_k(t)$ is convergent if

$$\int_0^t |\,d\{t\phi_k(t)\}\,| = O(t).$$

Digitized by Google

Now, if $k > 1$,

$$t\phi_k(t) = \frac{\Phi_k(t)}{t^{k-1}},$$

$$\frac{d\{t\phi_k(t)\}}{dt} = \frac{\Phi_{k-1}(t)}{t^{k-1}} - (k-1)\frac{\Phi_k(t)}{t^k},$$

and so

$$\begin{aligned}\int_0^t |d\{t\phi_k(t)\}| &\leqslant \int_0^t \frac{|\Phi_{k-1}(t)|}{t^{k-1}}\,dt + (k-1)\int_0^t \frac{|\Phi_k(t)|}{t^k}\,dt \\ &= t\phi_k^*(t) + (k-1)\int_0^t \phi_k(t)\,dt \\ &= t.O(1) + (k-1).o(t) = O(t),\end{aligned}$$

and Young's test is satisfied.

When $k = 1$

$$t\phi_1(t) = \Phi_1(t),$$

$$\int_0^t |d\{t\phi_1(t)\}| = \text{total variation of } \Phi_1(t) \text{ in } (0, t)$$

$$\begin{aligned}&= \int_0^t |\phi(t)|\,dt \\ &= t\phi_1^*(t) = O(t)\end{aligned}$$

as before.

The condition (10·1) is thus sufficient in all cases.

Digitized by Google

On the descriptive form of Taylor's theorem. By Mr S. POLLARD, Trinity College.

[*Received* 14 August, *read* 25 October 1926.]

1. The particular form of Taylor's theorem considered states that:

If $f(x)$ is n times differentiable at a then

$$(1\cdot1)\quad f(a+h)=f(a)+hf'(a)+\ldots+\frac{h^n}{n!}f^n(a)+o(\rho^n),$$

where $$\rho=|h|.$$

This may be termed the *descriptive* form of Taylor's theorem in that it gives the behaviour of the remainder but not its analytical expression. Its value, on account of the extremely compact hypothesis, can hardly be over-estimated.

The theorem was originally given by W. H. Young and rediscovered by Fowler, who devotes to it one of the appendices of his tract on Plane Curves. Both these writers obtain the Theorem from the theory of indeterminate forms.

A very casual inspection of the theorem will show that, if possible, it should be proved apart from the theory of indeterminate forms because it is the one result which, when established, simplifies this theory more than any other. The analysis about to be given shows that it can be deduced without much trouble from the ordinary form of Taylor's theorem due to Lagrange.

2. Write

$$F(h)=f(a+h)-f(a)-hf'(a)-\ldots-\frac{h^n}{n!}f^n(a).$$

Then (1·1) is the same as

$$\frac{F(h)}{h^n}\to 0$$

as $h\to 0$.

Now the successive derivatives of $F(h)$ up to the $(n-1)$th inclusive are respectively

$$F'(h)=f'(a+h)-f'(a)-\ldots-\frac{h^{n-1}}{(n-1)!}f^n(a),$$

$$F''(h)=f''(a+h)-f''(a)-\ldots-\frac{h^{n-2}}{(n-2)!}f^n(a),$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$F^{n-1}(h)=f^{n-1}(a+h)-f^{n-1}(a)-hf^n(a),$$

Digitized by Google

and these, since $f^n(a)$ exists, exist in some neighbourhood of $h=0$. Lagrange's form of Taylor's theorem can be applied to $F(h)$ and yields

$$F(h)=F(0)+hF'(0)+\ldots+\frac{h^{n-1}}{(n-1)!}F^{n-1}(\theta h) \quad (0<|\theta|<1)$$
$$=\frac{h^{n-1}}{(n-1)!}\{f^{n-1}(a+\theta h)-f^{n-1}(a)-\theta h f^n(a)\},$$

for all sufficiently small values of $|h|$.

Thus

$$\frac{F(h)}{h^n}=\frac{\theta}{(n-1)!}\left\{\frac{f^{n-1}(a+\theta h)-f^{n-1}(a)}{\theta h}-f^n(a)\right\}\to 0,$$

because $$\frac{f^{n-1}(a+\theta h)-f^{n-1}(a)}{\theta h}\to f^n(a).$$

3. It is possible, by a modification of the above argument, to obtain a more precise statement of the theorem. For let

$$\omega(\rho)$$

be the modulus of differentiability of $f^{n-1}(x)$ at a, i.e. the upper bound of

$$\left|\frac{f^{n-1}(a+t)-f^{n-1}(a)}{t}-f^n(a)\right|,$$

for $0<|t|<\rho$.

By the fractional form of Taylor's theorem, with Lagrange's remainder,

$$\frac{F(h)}{\frac{h^n}{n!}}=\frac{F^{n-1}(\theta h)}{\theta h}=\frac{f^{n-1}(a+\theta h)-f^{n-1}(a)}{\theta h}-f^n(a),$$

which is not greater in absolute value than $\omega(\rho)$.

Thus (1·1) can be put in the precise form

$$f(a+h)=f(a)+hf'(a)+\ldots+\frac{h^n}{n!}\{f^n(a)+\theta\omega(\rho)\},$$

where $|\theta|\leqslant 1$.

4. The theorem for a single variable has been extended by W. H. Young to the case of n variables. Here again a precise statement is possible. In the case of two variables it runs as follows:

If $f(x, y)$ is n times differentiable at (a, b) and $\Omega(\sigma)$ is the upper bound of the moduli of differentiability at (a, b) of the unction

$$f_{\mu,\nu}(x, y)=\frac{\partial^{n-1}}{\partial x^\mu \partial y^\nu}f(x, y) \qquad (\mu+\nu=n-1)$$

Digitized by Google

then

$$(4\cdot1)\quad f(a+h, b+k) = f(a, b) + df(a, b) + \ldots + \frac{1}{n!} d^n f(a, b) + \frac{\rho^n}{n!}\theta\Omega(\rho),$$

where $$\rho = |h| + |k|, \quad |\theta| \leqslant 1,$$

and $$d^r f(a, b) = \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^r f(x, y) \qquad (x = a, y = b).$$

The modulus of differentiability $\omega_{\mu,\nu}(\sigma)$ of $f_{\mu,\nu}(x, y)$ at (a, b) is the upper bound for $0 < |u| + |v| < \sigma$ of

$$\frac{1}{|u| + |v|}\left| f_{\mu,\nu}(a+u, b+v) - f_{\mu,\nu}(a, b) - u\frac{\partial}{\partial a} f_{\mu,\nu}(a, b) - v\frac{\partial}{\partial b} f_{\mu,\nu}(a, b)\right|,$$

and $\Omega(\sigma)$ is the greatest of the n numbers $\omega_{\mu,\nu}(\sigma)$ where $\mu + \nu = n - 1$.

Write $$\frac{h}{\rho} = l, \quad \frac{k}{\rho} = m,$$

so that $$|l| + |m| = 1,$$

and also $$\phi(t) = f(a + lt, b + mt),$$

so that $$f(a+h, b+k) = \phi(\rho).$$

It is known that $\phi(t)$ is n times differentiable at $t = 0$ with

$$\phi^r(0) = \left(l\frac{\partial}{\partial x} + m\frac{\partial}{\partial y}\right)^r f(x, y) \qquad (x = a, y = b)$$

$$= \frac{1}{\rho^r} d^r f(a, b).$$

Since $$f_{\mu,\nu}(x, y) = \frac{\partial^{n-1}}{\partial x^\mu \partial y^\nu} f(x, y) \qquad (\mu + \nu = n - 1),$$

it follows that

$$\frac{\phi^{n-1}(t) - \phi^{n-1}(0)}{t} - \phi^n(0) = \sum_{\mu+\nu=n-1} \frac{(n-1)!}{\mu!\,\nu!} \frac{l^\mu m^\nu}{t}$$

$$\times \left\{ f_{\mu,\nu}(a+lt, b+mt) - f_{\mu,\nu}(a, b) - lt\frac{\partial}{\partial a} f_{\mu,\nu}(a, b) - mt\frac{\partial}{\partial b} f_{\mu,\nu}(a, b)\right\},$$

so that, for $0 < t < \rho$

$$\left|\frac{\phi^{n-1}(t) - \phi^{n-1}(0)}{t} - \phi^n(0)\right| \leqslant \sum_{\mu+\nu=n-1} \frac{(n-1)!}{\mu!\,\nu!} |l|^\mu |m|^\nu . \Omega(\rho)$$

$$= \{|l| + |m|\}^{n-1}\Omega(\rho) = \Omega(\rho).$$

Thus, by the result of § 3,

$$\phi(\rho) = \phi(0) + \rho\phi'(0) + \ldots + \frac{\rho^n}{n!}\{\phi^n(0) + \theta\Omega(\rho)\},$$

which is plainly another way of writing (4·1).

Digitized by Google

The correspondence between lines in threefold space and points of a quadric fourfold in fivefold space, established by a geometrical construction. By Mr T. L. WREN, St John's College.

[*Received* 17 August, *read* 25 October 1926.]

1. This correspondence has been established by Mr H. W. Turnbull, *Proc. Camb. Phil. Soc.* vol. XXII. (1925), pp. 694—9, using a construction which leaves uncorrelated the lines of a special linear complex and the points of the quadric lying in one tangent fourfold; these have to be correlated by means of an independent construction.

The geometrical construction described below gives a direct representation of the lines of a threefold space by the points of a quadric fourfold in fivefold space, with the same exception; but it will be shown that this representation involves a definite one-one correspondence between the remaining lines and points for which the construction is not directly applicable.

For this construction we assume the existence of a quadric fourfold which has a definite tangent fourfold at each of its points, and contains three real planes with a common point, no two of which meet in a line. This assumption is suggested by the fact that the equation $p_{23}p_{14} + p_{31}p_{24} + p_{12}p_{34} = 0$, connecting the six Plücker coordinates of a line in threefold space, represents a quadric fourfold containing the three real planes $p_{23} = p_{31} = p_{12} = 0$, $p_{23} = p_{24} = p_{34} = 0$, and $p_{23} = p_{31} - p_{34} = p_{12} + p_{24} = 0$, any two of which meet only in the point $p_{23} = p_{31} = p_{12} = p_{24} = p_{34} = 0$. Starting from this assumption, the existence of the two systems of generating planes on the quadric fourfold will be established with the aid of the construction in question.

2. Consider, in fivefold space, a quadric fourfold F_4^2 containing three real planes, ρ, σ, τ, of which any two meet only in the point O. Then ρ, σ, τ must lie in T_4, the tangent fourfold at O.

Let S_4 be a flat fourfold not passing through O; and Ω_3 the threefold common to S_4 and T_4. Then Ω_3 cuts F_4^2 in a quadric κ_2^2 which has three real skew generators (lying in ρ, σ, τ respectively). The section of F_4^2 by T_4 is therefore a cone generated by the two systems of real planes which project the two reguli on κ_2^2 from O; two planes of the same system meet only at O, while two planes of different systems meet in a line through O.

3. Let a, b, c be three generators of the same regulus on κ_2^2; and π, π' two planes in S_4 which pass through c but do not lie in Ω_3. The planes π, π' determine a flat threefold Λ_3 belonging to S_4.

Digitized by Google

Any line l of Λ_3 which is skew to c will meet π, π' in distinct points P, P' not in Ω_3; the planes aP, bP' then meet in a single point M, in S_4 but not in Ω_3; and the line OM meets F_4^2 again in a single point L, not in T_4.

Conversely, any point L of F_4^2, not in T_4, is projected from O into a point M of S_4, not in Ω_3; the planes aM, bM meet π, π' respectively in distinct points P, P' which define a unique line l ($\equiv PP'$) not meeting c.

In this way *a one-one correspondence is established between all the real lines of* Λ_3 *which do not meet* c *and all the real points of* F_4^2 *which do not lie in* T_4. It remains to be proved that this correspondence involves a one-one correspondence between the lines of Λ_3 which meet c and the points of F_4^2 in T_4.

4. Consider now in Λ_3 a plane α which meets c in a single point E; and a point V of α not in π or π'.

As l describes the flat pencil through V in α, the planes aP, bP' describe projective axial pencils in two flat threefolds of S_4; these threefolds have in common a plane δ which must contain the locus of M.

The plane δ meets Ω_3 only in the line e common to the planes aE, bE; and cuts the pencils $a(P)$, $b(P')$ in two projective flat pencils in which e is self-corresponding. Hence the locus of M is a straight line x, which meets Ω_3 in a single point Z, lying on e.

If V is in π, but on c, the locus of M is a line x in the plane aV; and x still meets Ω_3 in a single point (ae): similarly for V in π'.

Thus in any case when l describes a flat pencil ϕ of which only one ray meets c, the locus of M is a line x meeting Ω_3 in a single point Z on a generator e of κ_2^2. The plane joining O to x cuts F_4^2 in a conic to which belongs the line OZ; the rest of this conic is therefore a line f on F_4^2, meeting T_4 in a single point on OZ. The line f is the locus of L.

The construction of paragraph 3 establishes a one-one correspondence between the rays of ϕ and the points of f, except for the ray meeting c and the point in T_4; these must therefore also correspond as members of the pencil ϕ and range on f, but it remains to be proved that all lines f which represent pencils ϕ containing a given line l, which meets c, must meet T_4 in the same point.

5. In Λ_3 take two fixed lines l_1, l_2 not meeting c, but having a common point V and lying in a plane α; then l_1, l_2 are represented on F_4^2 by points L_1, L_2 not in T_4.

Any flat pencil ϕ_1 to which l_1 belongs is represented on F_4^2 by a line f_1 through L_1; and any flat pencil ϕ_2 containing l_2 by a line f_2 through L_2.

Digitized by Google

First consider only pencils ϕ_1, ϕ_2 lying in α. Since any ϕ_1 has a ray not meeting c in common with all but one ϕ_2, the corresponding line f_1 must meet all but one f_2; and *vice versa*. It follows that all these lines f_1, f_2 must lie in a single plane β. Thus all the lines of the plane α, except those meeting c, are represented on F_4^2 by all the points of a plane β, except those in T_4; and there will be one such plane β for every plane α of Λ_3 which does not contain c.

Similarly, by considering only pencils ϕ_1, ϕ_2 having V for vertex, it may be proved that all the lines of Λ_3 through V, except those meeting c, are represented on F_4^2 by all the points of a plane ω, except those in T_4; and there will be one such plane ω for every point V of Λ_3 not on c.

6. Thus *on F_4^2 there is a system of real planes β representing all planes of Λ_3 not containing c; and a second system of planes ω representing all points of Λ_3 not on c.*

Two planes β will in general have just one common point outside T_4 (since the planes they represent have a common line not in general meeting c), and therefore in general two planes β cannot meet in a line. Similarly for two planes ω.

A plane β and a plane ω will in general have no common point outside T_4; but if they represent a plane α and a point V in α they will have a common line f, which represents the flat pencil through V in α.

Through any point L of F_4^2 not in T_4 there must pass a single infinity of planes β representing the planes through the line l represented by L, and a single infinity of planes ω representing the points of l. No two of these planes β can meet in any point besides L, and similarly for two planes ω; but each plane β will have a line through L common with any plane ω. All these planes β and ω must lie in the tangent fourfold at L; and they must form two systems analogous to the two systems of planes through O, since any flat fourfold not passing through L will cut them in generators of the two systems on a ruled quadric twofold.

The argument of paragraph 4 may be reversed to show that any line on F_4^2, not wholly in T_4, represents a flat pencil (with the usual exception); it then follows that any plane on F_4^2, not wholly in T_4, represents either a plane not through c or a point not on c. Hence *the planes β and ω through L are the only planes on F_4^2 through L.*

7. It is now evident that, in the construction of paragraph 3, the point O may be replaced by any point L of F_4^2 not in T_4. If L_0 be a point of F_4^2 in T_4 and we take L also outside the tangent fourfold at L_0, then L_0 is not in the tangent fourfold at L. Hence

Digitized by Google

through any point whatever of F_4^2 there will pass two systems of real planes on F_4^2, such that two planes of the same system have no other common point, while two of different systems have a line in common.

Thus *on F_4^2 there must be two distinct systems of real planes; two planes of the same system have always just one common point; two planes of different systems have in general no common point, but may in special cases have a common line.* One of these systems consists of all the planes β which represent planes not containing c, together with the planes of the same system in T_4; the other consists of all the planes ω representing points not on c, together with those of the same system in T_4.

All the planes of F_4^2 in T_4 must pass through O, since κ_2^2 is a proper ruled quadric.

8. Consider now the correspondence between the points of a plane ω and the lines of Λ_3 through the point V, not on c, represented by ω. Since all but one of the flat pencils through V are represented by lines f in ω, this correspondence must be a collineation.

In this collineation the pencil of planes (α) through a line l_0 meeting c must correspond to a flat pencil of lines (f) in ω; the vertex L_0 of (f) must be a point of T_4, since it cannot represent a line not meeting c.

All but one of the planes (α) through l_0 are represented by planes β on F_4^2; each of these planes β must contain the line f corresponding in the collineation to the associated plane α, and therefore passes through L_0.

Since two planes β can have but one common point, and L_0 is in the plane ω which represents V, it follows that L_0 must lie in all the planes ω which represent points of l_0.

Thus *the representation of lines of Λ_3 not meeting c by points of F_4^2 not in T_4 leads to a definite point L_0 of T_4 associated with a line l_0 meeting c;* namely the single point L_0 common to all the planes β, ω representing planes through l_0, and points on l_0, respectively.

Any point L_0 of F_4^2 in T_4, other than O, is the single point common to a set of planes β and a set of planes ω (paragraph 7); these represent respectively a set of planes α of which no two have a common line not meeting c, and a set of points V such that the join of any two meets c. From what has been said above it is evident that the line common to any two of these planes α must be common to all, and must coincide with the line joining any two of the points V.

Digitized by Google

The representation of lines not meeting c by points of F_4^2 not in T_4 therefore involves a one-one correspondence between the lines meeting c and the points in T_4 other than O; leaving the single line c to correspond to the single point O.

9. *The correspondence between lines meeting c and points in T_4 is of exactly the same character as the general correspondence.*

Any plane β meets T_4 in a line f_0 not passing through O; the points of f_0 represent the lines meeting c of the plane α represented by β.

Any plane ω meets T_4 in a line whose points represent the lines meeting c which pass through the point V represented by ω.

The planes β representing all planes α through a given point V_0 of c will cut T_4 in lines f_0, any two of which must meet in the point of T_4 corresponding to the common line of the two associated planes α; these lines f_0 therefore lie in one plane. Thus the lines through V_0 are represented by the points of a plane in T_4; this plane must belong to the ω-system since it is met by certain planes β in lines.

Similarly the lines of a plane through c are represented by the points of a plane in T_4 belonging to the β-system.

Thus the planes on F_4^2 which lie in T_4 and pass through O present no exception to the general representation; those of the β-system represent planes through c, while those of the ω-system represent points of c.

The character of the representation is therefore completely independent of the special elements O, S_4, a, b, c, π, π', chosen subject to the restrictions specified in paragraphs 2, 3; and these elements provide no singularities in the correspondence between the lines of Λ_3 and the points of F_4^2, although the construction fails to determine directly the point in T_4 corresponding to a line meeting c.

Digitized by Google

On some resistance properties of a certain network containing inductances and capacities and their analogies in a vibrating mechanical system. By Mr E. B. MOULLIN, Downing College. (Communicated by Dr G. F. C. SEARLE.)

[*Received* 1 May, *read* 26 July 1926.]

(1) *Introduction.*

In the network shown diagrammatically in Fig. 1, A_0A_3, A_3A_5, ... are resistances of values a_1, a_3, ... joined in series with one another, and A_3B_3, A_5B_5, ... are resistances of values $1/a_2$, $1/a_4$, ...: the points B_0B_3, ... are all on a cable of negligible resistance. The members A_0A_3, A_3A_5, ... will be called the series members of the network, and the members A_3B_3, A_5B_5, ... will be called the shunt members of the network: the points A_0B_0, will be called the input terminals of the network. If a potential difference is maintained between the input terminals, currents will flow in the members of the network.

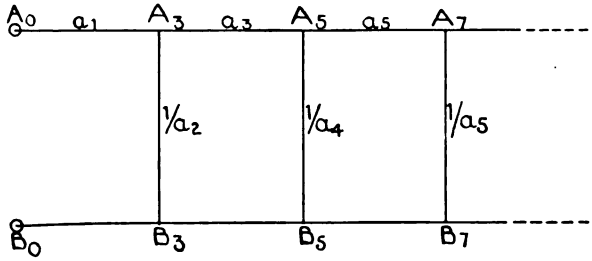


Fig. 1.

The more general network of which the system of Fig. 1 is a degenerate form is important in telegraphy and telephony; its properties depend on the character of the various members. In the degenerate case each member is a simple resistance, and then a_1, a_2, ... may represent portions of an overhead line, while $1/a_3$, $1/a_4$, ... represent leaks at the points of support. The special case when $a_1 = a_3$ ... and $1/a_2 = 1/a_4$ ... has been examined previously. (See Dr G. F. C. Searle, F.R.S., *Proc. Camb. Phil. Soc.* 1915, Vol. XVIII, p. 111, 'Calculations of the electrical resistance of a certain network of conductors.')

The resistance Z_0, measured between the input terminals of the network of Fig. 1 may be represented by the continued fraction

$$Z_0 = a_1 + \frac{1}{a_2+} \frac{1}{a_3+} \frac{1}{a_4+} \cdots \frac{1}{a_r+} \frac{1}{a_{r+1}+} \cdots \frac{1}{a_n} \quad \ldots\ldots(1).$$

We shall call Z_0 the input resistance of the network.

Digitized by Google

In an example of more general interest the branches of the network contain inductances and capacities, and an alternating potential difference $v = V \sin pt$ is maintained between the input terminals.

If a current $i = I \sin(pt + \alpha)$ then enters the network along A_0A_3 and leaves it along B_0B_3, we term the ratio V/I the impedance of the network; this quantity replaces the input resistance of the degenerate system. Important examples of the more general network are the "filter circuits" used in telephony. In these circuits the receiving apparatus is the final shunt member of the network and each series member combined with the succeeding shunt member is called a stage; thus a filter has as many stages as it has shunt members. The voltage between the input terminals is often not simple harmonic and it may be undesirable that currents of certain frequencies should flow through the receiver: a filter system is then used to reduce these undesired components of current to a negligible amount and the degree of elimination depends on the number of stages in the filter.

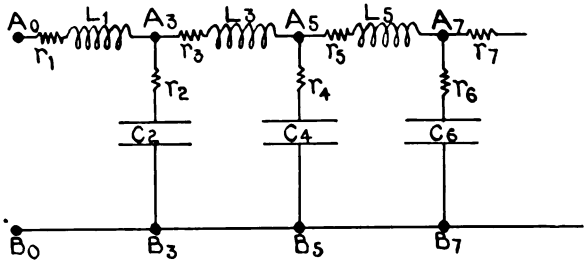


Fig. 2.

The circuit of Fig. 2 is called the "low pass" filter and is used to exclude from the receiver (for example A_7B_7) currents whose frequencies are greater than some predetermined value. If we transpose the inductances and the capacities, as in Fig. 3, we obtain a "high pass" filter which excludes from the receiver currents whose frequencies are less than a certain value. If the individual members of the network contain both inductance and capacity we may arrange to obtain a "band pass" filter or a "band stop" filter which permits or excludes respectively only currents whose frequencies lie within certain specified limits.

In connexion with these circuits, certain impedance problems arise which deserve attention in general. Thus if Z_0 is the input impedance of the network when the alternator is connected between the input terminals we may wish to know what the impedance will be at the same frequency if the input terminals are short circuited and the alternator is inserted in series with

Digitized by Google

some other shunt or series member. Again, since both inductance and capacity are present, resonance will be possible and we may enquire under what conditions resonance will occur. Further, if resonance occurs at a given frequency when the alternator is between A_0 and B_0, will resonance necessarily occur at the same frequency if A_0B_0 are joined and the alternator is inserted in some other member of the network? Also if resonance does occur again when the alternator is transferred to some other particular member, would it necessarily have occurred if the alternator had been transferred to any one arbitrarily selected member?

In either of the networks of Figs. 2 and 3, two forms of resonance are possible. In one case the system behaves as an "acceptor" circuit and the current it receives is a maximum and the impedance a minimum. In the other case the system behaves as a "rejector" circuit and the current is a minimum and the impedance a maximum. In the case in which the system is a rejector circuit to a given alternator placed in one member of the network, it is important to know if it will also be a rejector circuit when the

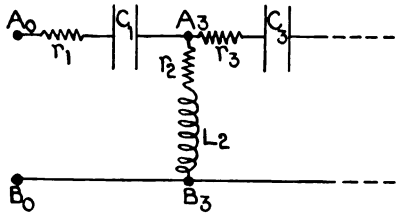


Fig. 3.

same alternator is placed in any one other member, or if it will then become an acceptor circuit or if it will fail to exhibit resonance.

Similar problems arise in mechanical vibrations of certain systems of interconnected masses, for the equations of the currents in the network of Fig. 2 are the same as the equations of motion for a system of fly-wheels fixed on a light elastic shaft turning freely in bearings, or of a suspended system of heavy particles connected together by a series of light vertical springs.

The torsional problem is of practical interest in the investigation of torsional oscillations of the crank shaft masses of a multi-cylinder engine, and here one of our problems is as follows:—If any one of the cranks is actuated by pressure behind a light piston and resonance of torsional oscillations is produced, will resonance be produced if the same piston actuates any other crank and what will be the result of actuating every crank simultaneously by a similar piston?

Digitized by Google

(2) *Expression for the impedance of the network.*

We may still use the continued fraction (1) to express the impedance of the general network, typified diagrammatically by Fig. 1 and exemplified in Figs. 2 and 3, if we specify the meanings to be attributed to the quantities a_1, a_2, ... The network is actuated by a periodic voltage $v = V\epsilon^{jpt}$ applied between A_0 and B_0, and in the final steady state, which we are here considering, the currents in the various branches will also be proportional to ϵ^{jpt}. We may therefore replace $\frac{d}{dt}$ by jp and $\frac{d^2}{dt^2}$ by $-p^2$: if this is done, we find that A_0A_3 in Fig. 3 is represented by the complex quantity $(r_1 + jpL_1)$. We may express the same fact more directly as follows:—Every current can be analysed into a Fourier series of simple harmonic terms and therefore every component may be represented instantaneously by the projection of a uniformly rotating line of constant length. Now the voltage across A_0A_3 consists of a component ri in phase with the current and a component pL_1i in phase quadrature with the current. So we may represent the voltage A_0A_3 by another uniformly rotating line which is related to the current line by the vector equation $V = I(r_1 + jpL_1)$. Hence, in accordance with the well-known methods of making alternating current calculations, we may replace the a's by complex quantities such that $a_1 = \{r_1 + jpL_1\}$ and $1/a_2 = \{r_2 - j/(pC_2)\}$, where $j \equiv \sqrt{-1}$.

We may now proceed to use the properties of continued fractions; the reader will note that all the quotients are in general complex quantities.

Let p_r/q_r be the rth convergent of fraction (1). Then

$$Z_0 = \frac{Y_r p_{r-1} + p_{r-2}}{Y_r q_{r-1} + q_{r-2}} = \frac{p_{r-1} + p_{r-2}/Y_r}{q_{r-1} + q_{r-2}/Y_r} \quad \text{.....(2),}$$

where
$$Y_r = a_r + \frac{1}{a_{r+1}} + \frac{1}{a_{r+2}} + \cdots \frac{1}{a_n}.$$

Now let A_0 and B_0 be joined and let the network be opened at A_r and let an alternator be inserted between A_r and A_r' as indicated in Fig. 4. We may note that the quantity Y_r defined above is the impedance of the part of the network to the right of the points A_r' and B_r'. Let X_r be the impedance of the network to the left of A_rB_r. Then according to the arrangement of Fig. 4 we have

$$X_r = \frac{1}{a_{r-1}} + \frac{1}{a_{r-2}} + \frac{1}{a_{r-3}} + \cdots \frac{1}{a_1}.$$

When the alternator is placed between the points A_r and A_r' the circuit consists of an impedance Y_r to the right of A_r' which

Digitized by Google

is in series with the impedance X_r to the left of A_r. Accordingly the impedance Z_r, between the points A_r and A_r' is given by the expression

$$Z_r = (X_r + Y_r) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3).$$

We wish to relate Z_0 and Z_r. It follows readily from the theory of continued fractions that

$$\frac{p_{r-1}}{p_{r-2}} = a_{r-1} + \frac{1}{a_{r-2}+} \frac{1}{a_{r-3}+} \cdots \frac{1}{a_1} = \frac{1}{X_r} \quad \ldots\ldots\ldots(4),$$

$$\frac{q_{r-1}}{q_{r-2}} = a_{r-1} + \frac{1}{a_{r-2}+} \frac{1}{a_{r-3}+} \cdots \frac{1}{a_2} \quad \ldots\ldots\ldots\ldots\ldots(5),$$

and

$$\frac{p_{r-1}}{p_{r-2}} - \frac{q_{r-1}}{q_{r-2}} = \frac{(-1)^{r-1}}{p_{r-2}q_{r-2}} \quad \ldots\ldots\ldots\ldots\ldots(6).$$

In our notation r is an odd integer. Hence by (4) and (6) we have

$$\frac{q_{r-1}}{q_{r-2}} = \frac{1}{X_r} - \frac{1}{p_{r-2}q_{r-2}}.$$

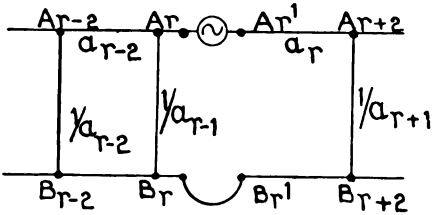


Fig. 4.

On substituting in (2), we have by (3) that

$$Z_0 = \frac{p_{r-2}\left(\frac{1}{X_r} + \frac{1}{Y_r}\right)}{q_{r-2}\left(\frac{1}{X_r} + \frac{1}{Y_r}\right) - \frac{1}{p_{r-2}}} = \frac{p_{r-2}^2 Z_r}{p_{r-2}q_{r-2}Z_r - X_r Y_r} \quad \ldots(7).$$

Equation (7) connects Z_0 with Z_r, X_r, Y_r, p_{r-2}, q_{r-2}.

(3) *Resonance conditions.*

We will now find the conditions for resonance in a network whose members are subject to the following restrictions.

(α) Every member is resistanceless, so that at resonance the impedance is zero or infinite according as the circuit is an acceptor or a rejector.

(β) No one of the impedances $a_1, a_3, a_5, \ldots$ is infinite and no one of the impedances $1/a_2, 1/a_4, 1/a_6, \ldots$ is zero. This means that

Digitized by Google

no series member may contain inductance and capacity in parallel and no shunt member may contain inductance and capacity in series.

Since p_r and q_r etc. consist of various products of the a's no p or q can be infinite on account of restriction (β).

(i) Acceptor circuit from input terminals.

We will first suppose that neither X_r nor Y_r is infinite. Then, from (7) Z_0 will be zero if $p_{r-2}=0$, or if $Z_r=0$, or if both p_{r-2} and Z_r are zero, provided that the denominator of (7) is not zero.

If $Z_r=0$, the system is an acceptor circuit, with A_0B_0 joined, to an alternator placed between A_r and A_r' and by (7) it is then also an acceptor circuit to an alternator placed between A_0 and B_0. Thus we have proved that the network can be an acceptor circuit when it is energised by an alternator either situated in series with the rth series member or connected between A_0 and B_0.

Now we see from (6) that $p_{r-1}q_{r-2}-p_{r-2}q_{r-1}=1$, since r is odd. Hence, if p_{r-2} is zero, then p_{r-1} is not zero, since no p or q is infinite.

But, by (4), we have $p_{r-1}/p_{r-2}=1/X_r$ and hence, if $p_{r-2}=0$, then $X_r=0$ also.

But we have seen from (7) that Z_0 is zero if $p_{r-2}=0$ and we see from (2) that this entails that $Y_r=0$, since neither p_{r-1} nor q_{r-2} can be zero when $p_{r-2}=0$; accordingly $X_r=0$ and $Y_r=0$ and $Z_r=(X_r+Y_r)=0$. Hence the necessary and sufficient condition for the case where X_r and Y_r are finite is that Z_0 should be zero. Thus we have proved that if the network is an acceptor circuit to an alternator placed in series with a certain series member then it must be an acceptor circuit to an alternator placed between the input terminals.

If one and one only of X_r and Y_r is infinite then $Z_r=X_r+Y_r$ will certainly be infinite; is it then possible that Z_0 should be zero?

We see from (2) that if Y_r is infinite, then Z_0 is zero only if p_{r-1} is zero. But we see from (4) that if p_{r-1} is zero then X_r is infinite. Hence we find that Z_0 is zero if $X_r+Y_r=0$ or if both X_r and Y_r are infinite*. In the first case the system is an acceptor from both positions of the driving alternator and in the second condition it is an acceptor from one place and a rejector from the other.

* Though both X_r and Y_r become infinite at a certain frequency, it may be shown that their sum cannot be zero. Both X_r and Y_r consist of two parallel branches, one whose impedance is positive and increasing and the other whose impedance is negative and decreasing for a steady increase of frequency: hence the rate of change of each branch impedance is positive. From this it follows that X_r/Y_r cannot have the value -1.

Digitized by Google

We have now proved that if the network is an acceptor circuit from the input terminals, then the same alternator will produce resonance if it is inserted in any series member whatsoever provided the input terminals have been short-circuited. There will always be resonance but not necessarily acceptor resonance. There will be rejector resonance if, and only if, the alternator is placed in a member which previously carried no current: that is to say only if it is situated in a member which carried no current, and so may be called a current node, when the alternator was feeding the input terminals. Conversely we see that if an alternator at the input terminals causes a certain member to be a current node, then this member will remain a current node if the alternator is moved to it.

(ii) Rejector circuit from input terminals.

We will now examine the condition that the network should be a rejector circuit from the input terminals, in which condition Z_0 must be infinite. Remembering that no p or q is infinite we see by equation (2) that Z_0 is infinite only if

$$Y_r = -q_{r-2}/q_{r-1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (8).$$

Now both Z_r and Z_0 can be infinite only if either X_r or Y_r is infinite: if both are infinite we have seen above that Z_0 is zero. Now if q_{r-1} is zero, neither p_{r-1} nor q_{r-2} can be zero. Therefore both Z_0 and Z_r will be infinite if Y_r is infinite and q_{r-1} is zero simultaneously: also they will both be infinite if X_r is infinite and Y_r satisfies condition (8). We may note that if X_r is infinite q_{r-2} cannot be zero and so $Y_r = 0$ cannot satisfy (8), but that if Y_r is infinite the condition that $q_{r-1} = 0$ does not preclude the possibility that X_r should be zero.

So we find that if the network is a rejector circuit from the input terminals it will not be a rejector circuit when energised from some series member unless certain special conditions are satisfied.

(4) *Network excited from a shunt member.*

We will now consider briefly the conditions obtaining when the alternator is placed in a shunt member as shown in Fig. 5. Let X_r' be the impedance of the part of the network to the left of A_rB_r (excluding the shunt member A_rB_r). Then the impedance Z_r' presented to the alternator, consists of the impedance X_r' in parallel with the impedance Y_r, and these two together are in series with the impedance $1/a_{r-1}$. Hence

$$Z_r' = \frac{X_r' Y_r}{X_r' + Y_r} + \frac{1}{a_{r-1}}.$$

Digitized by Google

Now $$X_r' = a_{r-2} + \frac{1}{a_{r-3}} + \frac{1}{a_{r-4}} + \cdots \frac{1}{a_1},$$

and then $$X_r = \frac{1}{a_{r-1} + 1/X_r'}.$$

Hence $$X_r' = -\frac{X_r}{(X_r a_{r-1} - 1)} \quad \ldots\ldots\ldots\ldots\ldots\ldots (9).$$

Now $$Z_r' = \frac{X_r'(Y_r a_{r-1} + 1) + Y_r}{a_{r-1}(X_r' + Y_r)} = \frac{a_{r-1} + 1/Y_r + 1/X_r'}{a_{r-1}(1/Y_r + 1/X_r')} \ldots (9)a,$$

whence, using (9) we have

$$Z_r' = \frac{1/X_r + 1/Y_r}{a_{r-1}(1/Y_r + 1/X_r - a_{r-1})} \quad \ldots\ldots\ldots\ldots\ldots\ldots (10).$$

We see from (10) that $Z_r' = 0$ if $X_r + Y_r = 0$, or if $X_r + Y_r$ are both infinite. But we saw previously that $Z_0 = 0$ in both these circumstances. If therefore the circuit is an acceptor circuit from

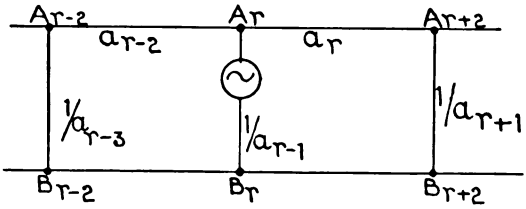


Fig. 5.

the input terminals it will always be an acceptor circuit to an alternator placed in any shunt member, provided only the impedance of that particular member is not infinite.

The conditions that the circuit of Fig. 5 should be a rejector circuit are readily followed from (9)a. In order that the impedances Z_0 and Z_r' should be infinite simultaneously, very special selection of the members of the network may be necessary.

If we remove restriction (β) we can examine the resonance conditions only when we have specified which of the a's have become infinite.

(5) *Illustration.*

To illustrate the various resonance conditions just developed let us consider the circuit depicted in Fig. 6, which represents a five-stage filter in which all the inductances are equal to L and all the capacities are equal to C and all the resistances are zero. If we write $1/LC = p_0^2$, and make suitable substitutions in (1) we can

Digitized by Google

show that the system is an acceptor circuit to an alternator placed between A_0 and B_0 when p^2/p_0^2 is a root of the equation

$$y^6 - 10y^4 + 36y^3 - 56y^2 + 35y - 6 = 0,$$

and the roots of this are

$$y = 0{\cdot}27,\ 1,\ 2,\ 3,\ 3{\cdot}73.$$

Acceptor condition for $p^2/p_0^2 = 1$.

That $p = p_0$ is a possible condition is readily seen, for at this frequency the portion to the right of A_9B_9 has infinite impedance and its presence is not felt by C_4. But L_4C_4 is an acceptor circuit and therefore is a complete short circuit across C_3 which will therefore carry no current. Hence A_7 is in effect joined to B_7 and hence the portion to the right of A_3 has infinite impedance and in effect we are left with only L_1 and C_1 in series. The current will be large everywhere except in L_2, L_5 and C_3 where it is zero.

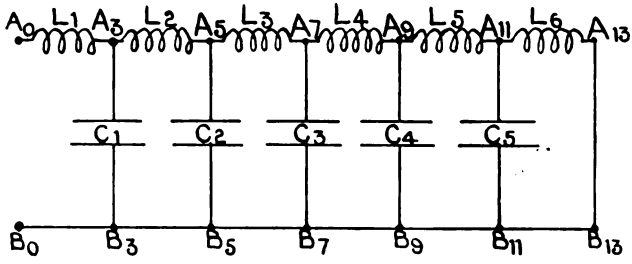


Fig. 6.

Let us now join A_0B_0 and place an alternator in series with L_4. The portion to the right of A_7 has zero impedance and thus $Y = 0$. The portion to the left of C_3 reduces to a zero impedance shunting C_3 and hence $X = 0$ and the condition $X + Y = 0$ is satisfied.

If we now remove the alternator from L_4 and place it in series with L_5, we see readily that Y is infinite. But we have just seen that C_3 is short-circuited by the system to its left and so in effect A_7 is joined to B_7. Consequently X consists of L_4 in parallel with C_4 and this circuit has infinite impedance. Consequently the condition that both X and Y are infinite is satisfied and we have an acceptor circuit.

Acceptor condition for $p^2/p_0^2 = 2$.

Now consider the condition when $p^2 = 2/LC$, corresponding to the root $p^2/p_0^2 = 2$: then each inductance has an impedance $pL\sqrt{2}$ and each condenser has an impedance $1/(pC\sqrt{2})$. If we denote the inductive impedance by $2R$, then the condensive impedance will

Digitized by Google

equal $-R$. Then C_5 and L_6 in parallel evidently have impedance $-2R$ and this in series with L_5, whose impedance is $2R$, will make a short circuit across C_4. Following the same argument, we see that C_2 is short-circuited and that the initial part of the system (L_1, C_1, L_2) has zero impedance and thus the current will be large everywhere except in C_2 and C_4 where it is zero.

Now let us insert an alternator at A_7 in series with L_4. We have just seen that C_4 is short-circuited by the acceptor circuit beyond it and so Y reduces in effect to L_4 alone, and this impedance we have called $2R$. We now consider the circuit to the left of A_7B_7, and see that C_2 is short-circuited by the acceptor circuit beyond it. Therefore in effect X consists of L_3 in parallel with C_3 and evidently the impedance of this is $-2R$. Therefore we find $X=-2R$, and $Y=2R$, and so the condition $X+Y=0$ is satisfied though neither X nor Y is zero: the system is an acceptor circuit whether the alternator is at A_0 or A_7.

Rejector conditions with $p^2/p_0^2=1$.

To consider the rejector circuit conditions we will suppose that in Fig. 6, $A_{11}B_{11}$ are joined by a thick conductor. We then have a system which is a rejector circuit to an alternator placed between A_0 and B_0, when $p=p_0$. For the circuit beyond A_7B_7 evidently has infinite impedance and does not affect C_3: again, C_2 is short-circuited by L_3 in series with C_3 and evidently the remaining circuit has infinite impedance. Now let us insert the alternator at A_7 (and remember that $A_{11}B_{11}$ are permanently joined). Inspection shows that Y is infinite and that X is zero. But we saw that if Y is infinite then q_{r-1} must be zero if Z_0 is to be infinite. Now consider the circuit from A_0B_0 to A_7B_7 and suppose the remainder removed. Then C_2 is short-circuited by L_3, C_3 and so the impedance between A_0 and B_0 is infinite: this means (since $r=7$) that q_6 is zero and so the necessary condition is satisfied. We may note that the system up to A_7B_7 is a rejector circuit from one end and an acceptor circuit from the other. If the alternator is inserted at A_5, inspection will show that $Y=0$ and that $X=1/(p_0C)$: but if $Y=0$ the condition that Z_0 should be infinite is that $q_{r-2}=0$. Consequently (since now $r=5$) q_3 should be zero, and inspection shows that this is so.

(6) *Analogy with torsional oscillations.*

It has been stated previously that the equations of motion for a system of fly-wheels fixed to an elastic shaft are the same as the equations for the currents in the network of Fig. 2. If the comparison is made, it will be found that the twist in any portion of the shaft corresponds to the current in the corresponding condenser and that the displacement of any fly-wheel corresponds to the

Digitized by Google

current in the corresponding inductance. The five-stage filter illustrated in Fig. 6 is analogous to six equal fly-wheels fixed at even distances along a light uniform shaft which is mounted freely in bearings. This is illustrated in Fig. 7.

When considering Fig. 6 we saw that when $p = p_0$, the current in C_3 was zero and likewise the currents in L_2 and L_5, and that the current in L_3 equals the current in L_4 in both magnitude and phase. Now the current in L_6 lags a quarter period behind the P.D. between A_{11} and B_{11}, and hence behind the P.D. between A_9 and B_9: but the current through C_4 leads this P.D. by a quarter period and so the currents in L_4 and L_6 are in antiphase with one another. Similarly the currents in L_1 and L_3 are in antiphase.

Now let us translate these results to the system of Fig. 7 which we shall suppose to have been set in motion by applying an alternating couple, of frequency $p_0 = 1\Big/\sqrt{\frac{I_1}{\lambda_1}}$ for a short time only to the wheel of moment of inertia I_1, fixed to a shaft of strength λ_1. If the couple is then removed, each wheel of this frictionless

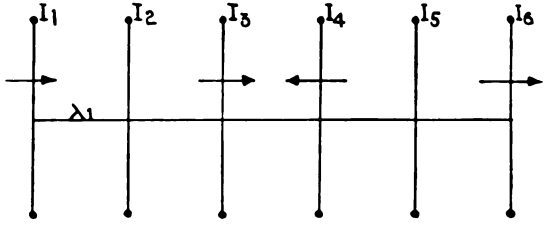


Fig. 7.

system will continue to oscillate at the frequency p_0. The system is exactly analogous to Fig. 6 and we can find the mode of vibration from the information we now have about the currents in Fig. 6.

Thus, since the current in C_3 is zero the angle of twist of the middle section of shaft must also be zero and therefore I_3 and I_4 must oscillate always as one body. Again the current in L_2 and L_5 is zero and therefore I_2 and I_5 should be at rest. Further the currents in L_1 and L_6 were in phase with one another and in antiphase to the current in L_3 and L_4, hence I_1 and I_6 should oscillate in phase with one another but in antiphase with I_3 and I_4. Arrowheads are marked in Fig. 7 to suggest this condition of motion. Evidently this motion is possible, for it gives zero angular momentum to the system and could have been initiated, for example, by twisting I_1 and I_6 through a clockwise angle θ and I_3 with I_4 through an anticlockwise angle θ and then releasing the system. We should then have the familiar two node vibration with nodes at I_2 and I_5.

Digitized by Google

Now consider the case when $p^2 = 2p_0^2$ in Fig. 6. The current through C_3 is twice as large as that through L_6 and consequently the currents through L_5 and L_6 must be in antiphase with one another. The current in C_4 is zero so that the current in L_4 and L_5 is the same in magnitude and phase and equal to half the current through C_3. Proceeding to the beginning of the circuit we find the currents are arranged symmetrically about the middle member. Now translating this to Fig. 7 we find I_2 and I_3 oscillating as one body, likewise I_4 and I_5. The displacement of I_2, I_3 and I_6 are always identical and in antiphase with I_1, I_4 and I_5. Again the motion is a possible one for a free system as the total angular momentum is zero: it is a three node vibration with nodes at the middle of the first, third and fifth section of the shaft.

Lastly consider the case when $p^2 = 3p_0^2$ in Fig. 6. The current in L_5 is twice the current in L_6 and is in antiphase with it whilst the current in C_3 is zero. If we proceed in a similar manner to consider all the members of the network, we find the currents distributed symmetrically on either side of the middle condenser. Translating this to Fig. 7, we find I_3 and I_4 moving as one body and with amplitude θ and that I_1 and I_6 move in phase with them also with amplitude θ. But I_2 and I_5 move with amplitude 2θ, and in antiphase with the other fly-wheels. This is again a possible condition for a free motion and gives the four node vibration with nodes one-third the way along the first and fifth section of the shaft (each distance measured from the terminal mass) and again at two-thirds along the second and fourth section (distances measured from the second and from the fifth mass). The single node and the five node vibrations are followed less readily because of the non-integral values of p^2.

Digitized by Google

The Polarisabilities of Atomic Cores. By BERTHA SWIRLES, Yarrow Research Student, Girton College. (Communicated by Mr. R. H. FOWLER.)

[*Received* 1 July, *read* 26 July 1926.]

1. *Introduction.* If it is assumed that the series electron of an atom polarises the core, then it has been shown by Born and Heisenberg* that the polarisability α of the core in a given state may be calculated from the corresponding term value by means of the approximate formulae,

$$\alpha = a_1^3 \cdot \frac{4n^2k^5}{3n^2 - k^2} \cdot q \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1),$$

$$\alpha = \frac{a_1^3 \cdot 2n^5k^5}{R \cdot (3n^2 - k^2)} \cdot \Delta\nu \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1'),$$

where

q is the quantum defect,

$\Delta\nu$ is the difference between the term and the corresponding hydrogen term,

R is the Rydberg constant in cm.$^{-1}$,

a_1 is the radius of first hydrogen orbit,

n is the principal quantum number,

k is the azimuthal „ „

For terms with small quantum defect either of the formulae (1) and (1′) may be used, but for terms with large quantum defect (1) gives a higher degree of accuracy.

Schrödinger in a recent paper† has suggested that there is a dispersion formula connecting the polarisability and the frequency of the polarising electron. Here "frequency" does not mean the frequency of rotation of the electron in its orbit, but its frequencies of transition to other orbits. In general it is necessary to consider only one—the strongest—of these transition frequencies.

The formula used by Schrödinger therefore resembles that obtained by Kramers and Heisenberg in their theory of dispersion‡. They obtain for the electrical moment M of an atom in its normal state on which there is incident radiation of frequency ν, the expression

$$M = \frac{E \cdot e^2}{4\pi^2 m} \sum_a \frac{f_a}{\nu_a^2 - \nu^2} \cos 2\pi\nu t \quad \ldots\ldots\ldots\ldots\ldots\ldots(2),$$

where $E \cos 2\pi\nu t$ is the electric force at time t, e and m have their usual meanings, the ν_a's are the absorption frequencies of the atom, and the f_a's are the strengths of the corresponding transitions. We

* M. Born u. W. Heisenberg, *Zeitschr. f. Phys.* 23, p. 388, 1924.

† E. Schrödinger, *Ann. d. Phys.* (4) 77, p. 43, 1925.

‡ Kramers u. Heisenberg, *Zeitschr. f. Phys.* 31, p. 681.

Digitized by Google

have also $f = a\tau_\nu$, where $\tau_\nu = 3c^3 m/(8\pi^2 e^2 \nu^2)$, and a is the Einstein probability coefficient. If in this formula we put $M/(E\cos 2\pi\nu t) = \alpha$, the ν_a's equal to the absorption frequencies of the core and ν equal to the principal transition frequency of the series electron, we have

$$\alpha = \frac{e^2}{4\pi^2 mc^2} \sum_a \frac{f_a}{\nu_a^2 - \nu^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3),$$

where the frequencies are now measured in cm.$^{-1}$.

The expressions (1) and (3) for α are obtained from entirely different considerations. To obtain the equations (1), (1′) it is assumed that the series electron occupies an orbit with quantum numbers n and k, such that the radius vector r joining it to the centre of the core is sufficiently large for the field to be treated as homogeneous in the region of the core and equal to e/r^2. The expression (3) is obtained by the methods of the Correspondence principle and not directly from a mechanical model. The two expressions are in a sense supplementary, and it will be assumed that they may be equated. A fuller justification may be possible with the aid of the new Quantum Mechanics. The theory has been applied to some elements with one series electron, for which the necessary data were available, in particular the elements of the Aluminium group and ionised Silicon, and Sodium and the corresponding "Stripped atoms."

2. *The polarisability of the Aluminium core.*

For convenience, a quantity $\alpha' = 10^5 Ra/a_1^3$ has been calculated instead of α itself. "Half quantum numbers" have been used for k. The following tables show the values of α' calculated from the d, f and g terms of the AlI spectrum.

TABLE 1. Al I d-terms*.

n	ν_n cm.$^{-1}$	q	$3p_1 - nd_1$	α'
3_1†	15842·0	·3681	32324·7	68·37
3_2	9345·1	−·4267	38821·6	−79·27
4	6041·1	−·2620	42125·6	−43·02
5	4109·9	−·1673	44056·8	−26·06
6	2932·9	−·1167	45233·8	−17·72
7	2183·6	−·0890	45983·1	−13·28
8	1682·1	−·0769	46484·6	−11·36
9	1334·7	−·0673	46832·0	− 9·87

* The term values used were calculated by Mr Hartree from lines given by A. Fowler and the third line of the f-series given by Meissner; $5g$ from a probable $4f$–$5g$ line given by Paschen-Götze.

† The occurrence of two terms with principal quantum number 3 needs some explanation. The unperturbed third orbit, i.e. with term value 12192·8 cm.$^{-1}$ (see Table 4), has a principal transition frequency 35973·9 cm.$^{-1}$, which lies very near a resonance frequency of the core. The presence of the core perturbs this orbit and may make it either more or less closely bound; in one case the corresponding

Digitized by Google

TABLE 2. Al I *f*-terms.

n	ν_n cm^{-1}	$\Delta\nu$	$3d_1 - nf$	α'
4	6960·1	101·7	8881·9	30·61
5	4449·3	59·9	11392·7	31·35
6	3085·9	37·7	12756·1	32·19

TABLE 3. Al I *g*-term.

n	ν_n cm.$^{-1}$	$\Delta\nu$	$4f - ng$	α'
5	4403·8	13·6	2556·3	28·65

The occurrence of negative q's in Table 1 requires some explanation. Table 4 shows a comparison of the Al *d*-terms with the corresponding hydrogen terms.

TABLE 4.

n	3	4	5	6	7	8	9
R_{Al}/n^2	12192·8	6858·4	4389·4	3048·2	2239·5	1714·6	1354·8
d-term	15842·0	9345·1	6041·1	4109·9	2932·9	2183·6	1682·1
q	·3681	·5733	·7380	·8326	·8832	·9110	·9231

From this comparison it appears that the term-differences and quantum defects are large, and that the terms are less hydrogen-like than might be expected. It seems probable that for comparison the 4th *d*-term and the succeeding ones should be moved one place to the left. In Table 1 this has in fact been done.

Referring to equation (3) we have

$$\alpha' = \frac{R}{a_1^3} \cdot 10^{-5} \cdot \frac{e^2}{4\pi^2 m} \cdot \frac{1}{c^2} \Sigma \frac{f_a}{\nu_a^2 - \nu^2}$$

$$= 5{\cdot}196 \cdot 10^{10} \Sigma \frac{f_a}{\nu_a^2 - \nu^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots (4).$$

The ν_a's may be expected to lie near the frequencies of the resonance lines of the core, but not exactly to coincide with them owing to the presence of the series electron.

transition frequency may be above, in the other below the resonance frequency of the core. If we assume that the smaller transition frequency corresponds to the more closely bound orbit, we can see that both orbits are possible; for an electron in the orbit with smaller transition frequency polarises the core positively so that the orbit is in fact more closely bound, while an electron in the orbit with greater transition frequency polarises the core negatively, so that the core repels it and the orbit is less closely bound. See p. 58 of Schrödinger's paper quoted above.

Digitized by Google

The resonance lines of the Al II spectrum are

$3S-3P$ $\nu = 59845{\cdot}1$ observed
$3S-3p_2$ $\nu = 32776{\cdot}6$ „
$3S-4P$ $\nu = 104300$ $4P$ term estimated
$3S-5P$ $\nu = 122000$ $5P$ „ „
$3S-4p_2$ $\nu = 100761$ from observed $4p$ term

Of these the first two lie in or near the range of frequencies in the fourth column of Tables 1, 2 and 3. There is little information obtainable with regard to the f's. Reiche and Thomas* in a recent paper give a summation rule for them, which is obtained by a use of the Correspondence principle. For a system containing one series electron the polarisation, in the classical theory, is expanded in a Fourier series, and by means of the relation between the coefficients in this series and the Einstein probability coefficients, the rule

$$\Sigma f_a = \frac{s}{3}$$

is obtained for a non-degenerate system. Here s is the degree of periodicity of the system and the summation is taken over all the lines of the series. For a degenerate system weight factors have to be introduced, but as in this case they also occur in equation (3) it may be assumed that they are included in the f_a's and the result remains unchanged. Four quantum numbers are required to describe the motion of a single outer electron in the atom, and in view of the suggestion by Goudsmit and Uhlenbeck that the electron spins about an axis, it appears that s should be taken equal to 4.

For a system with two or more outer electrons the uniformising variables for the motion in the classical theory cannot be found, and hence the expansion of the polarisation in a Fourier series is unknown. If, however, it is assumed that such an expansion is possible, then the analysis for one outer electron may be applied, and again

$$\Sigma f_a = \frac{s}{3},$$

where now $s = 4p$, where p is the number of outer electrons.

For Al^+, $p = 2$, and therefore

$$\Sigma f_a = 2\tfrac{2}{3}.$$

Paschen† gives the intensities of the lines $3S-3P$ and $3S-3p_2$ in the ratio 20 : 8. It therefore seems reasonable to take $f_a = 1$ approx. for $3S-3P$ and $f_a = {\cdot}4$ approx. for $3S-3p_2$.

If the resonance lines were unchanged by the presence of the series electron this would give a formula

* Reiche u. Thomas, *Zeitschr. f. Phys.* 34, p. 510.
† Paschen, *Ann. d. Phys.* 71, p. 537.

Digitized by Google

$$\alpha' = 5{\cdot}196 \,.\, 10^{10} \left\{ \frac{1}{59845{\cdot}1^2 - \nu^2} + \frac{{\cdot}4}{32776{\cdot}6^2 - \nu^2} \right\}$$

$$= \frac{500}{35{\cdot}81 - \frac{\nu^2}{10^8}} + \frac{200}{10{\cdot}74 - \frac{\nu^2}{10^8}} \text{ approx.} \quad \ldots\ldots\ldots\ldots\ldots\ldots (5).$$

If this curve is plotted it is seen to be similar in shape to that obtained from the last two columns of Tables 1, 2 and 3, but the agreement is not good.

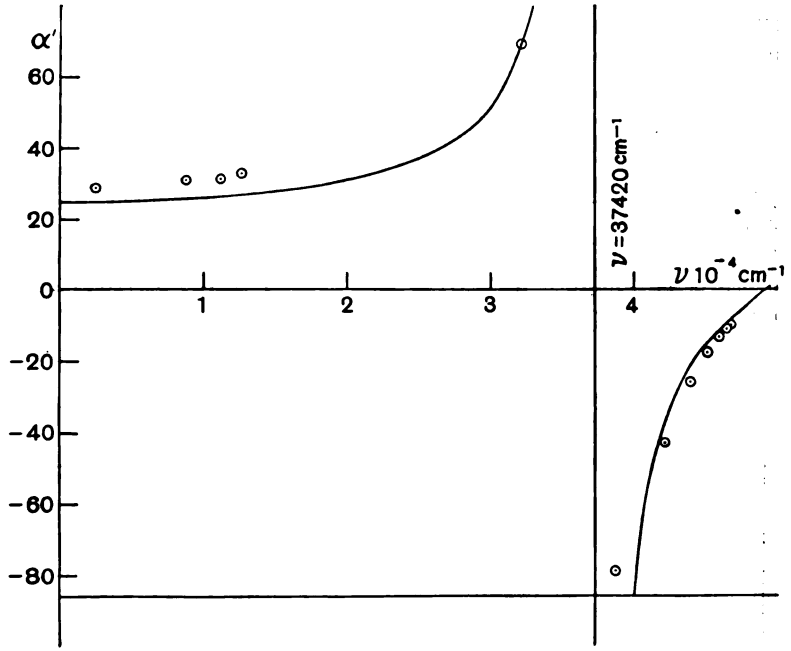


Fig. 1. Polarisability of Al I core.

If however

70710 cm.$^{-1}$ be substituted for 59845·1 cm.$^{-1}$

and 37420 cm.$^{-1}$ „ „ „ 32776·6 cm.$^{-1}$,

and in consequence the formula altered to

$$\alpha' = \frac{500}{50 - \frac{\nu^2}{10^8}} + \frac{200}{14 - \frac{\nu^2}{10^8}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (6),$$

much better agreement is obtained. See Fig. 1.

Digitized by Google

The polarisabilities were also calculated for the d-terms on the assumption that the q's remain positive throughout, and it was found that agreement could not be obtained by any reasonable modification of equation (5).

The change assumed in the resonance lines appears at first sight to be unreasonably large. A comparison, however, with the quadratic Stark effect of the alkalis, taking the field in the region of the core due to the electron to be e/r^2, where r is of the order of 10^{-8} cms., shows that it is of the order to be expected, for an alkali-like core, but of the opposite sign. The Al I core is, however, to be compared with the atom of an alkaline earth, and until the Stark effect of the alkaline earths has been more fully studied, a more definite comparison is not possible.

3. *Other elements of the Aluminium group and Ionised Silicon.* For Gallium and Indium only the d-lines are known with any certainty, and although the chief resonance line for each core can be fairly reliably identified, the line corresponding to the Al $3S - 3p_2$ line cannot. An equation can be guessed which gives fairly good agreement with the calculated polarisabilities, but for these reasons it is not of much value. It is uncertain whether the d-orbits of Thallium penetrate the core; if they are treated as non-penetrating orbits it is certainly impossible to explain the calculated polarisabilities by means of a dispersion formula.

For Si II more information is available. The following tables show the values of α' calculated from the d and f-terms:

TABLE 5. Si II d-terms*.

n	ν_n cm.$^{-1}$	q	$2p_1 - nd_1$	α'
3	52466·5	·1083	79064·5	5·03
4	30799·1	·2258	100731·9	9·27
5	19428·8	·2481	112102·2	9·67
6	13301·4	·2569	118229·6	9·74

TABLE 6. Si II f-terms.

n	ν_n cm.$^{-1}$	q	$3d - nf$	α'
4	28265·4	·060	24201·1	15·55
5	18061·4	·071	34405·1	16·43
6	12510·4	·078	39956·1	16·95
7	9169 } unreliable	·082	43297·5	17·20
8	7004 } unreliable	·084	45462·5	17·25
9	5524 } unreliable	·087	46942·5	19·30

* All Si terms are taken from a paper by A. Fowler, *Phil. Trans.* ccxxv, p. 1.

Digitized by Google

The resonance lines of Si III which lie in or near the range of frequencies are

$$3S - 3p = 82857,$$
$$3S - 3p_2 = 39330.$$

If the corresponding f's are taken to be 1·3 and ·3 approximately, and the lines are displaced to 137800 and 54770, so that the formula (4) becomes

$$\alpha' = \frac{700}{190 - \frac{\nu^2}{10^8}} + \frac{160}{30 - \frac{\nu^2}{10^8}},$$

fairly good agreement with the calculated values is obtained. See Fig. 2.

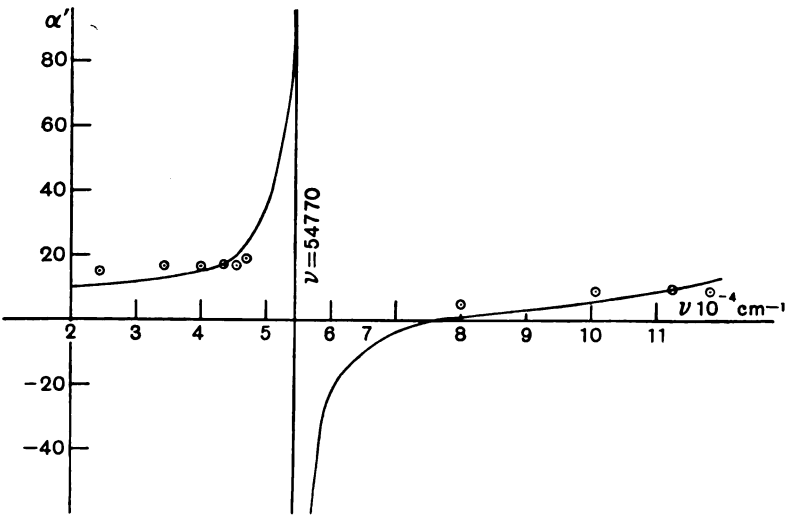


Fig. 2. Polarisability of Si II core.

4. *Sodium and the corresponding "stripped atoms."*

The polarisabilities of the cores deduced from the spectre of Na I, Mg II and Al III are much smaller than those previously considered. They are also constant approximately within a given series, and the variation for the different series is small. The frequencies of the resonance lines may be taken to be those of the lines discovered by Millikan and Bowen* and considered by them to belong to the X-ray spectrum. By analogy with the Neon spectrum, the interpretation as resonance lines of the core is more probably correct. These are in each case large compared with the frequencies of the series electron and to a sufficient degree of accuracy ν^2 may be omitted in equation (4). The mean value of α' for each series of terms is given in Table 7, together

* R. A. Millikan and I. S. Bowen, *Phys. Rev.* I. 1. 1924.

Digitized by Google

with the calculated values from equation (4). Since there are six outer electrons

$$\Sigma f_a = \frac{4.6}{3} = 8.$$

If f_a be taken equal to anything between 3 and 6 for the particular resonance line considered, fairly good agreement is obtained. The data are not sufficiently accurate to decide the value of f_a more certainly. The change in ν_a due to the presence of the series electron would in these cases probably be small compared with ν_a and need not be taken into account.

TABLE 7.

Element	α'_d	α'_f	α'_g	$\nu_a . 10^{-5}$ cm.$^{-1}$	$\alpha' = \frac{5.20.10^{10} \times f_a}{\nu_a^2}$	
					$f_a = 3$	$f_a = 6$
Na I	2·15	—	—	2·66	2·16	4·32
Mg II	1·57	·68	·51	4·32	·84	1·68
Al III	1·16	·51	·79	6·16	·42	·84

Note on a paper, "Über die Termgrossen der Serienspektren," *Zeitschr. f. Phys.* XXXVI. p. 92, by Albrecht Unsöld, which appeared during the course of the foregoing work.

In this paper a formula

$$\alpha = \frac{pe^2}{(2\pi c)^2 . m\nu_0^2}$$

is given for the polarisability of the core by the series electron. ν_0 is the chief resonance-frequency of the core, and the frequencies of the series electron are neglected throughout; this procedure is not always justifiable, particularly in the case of the Aluminium group. For these elements the equation certainly gives good agreement for the *f*-terms, but calculation shows that it is quite inaccurate for the *d*-terms.

Moreover, p is the number of loosely bound electrons in the core. The assumption is made that:

"A core, with p loosely bound electrons and the resonance frequency ν_0, behaves in its interaction with an outer electron as an aggregate of p isotropic harmonic oscillators of frequency ν_0."

In the dispersion theory of Kramers and Heisenberg p is replaced by a probability coefficient f. Reiche and Thomas give the rule that the sum of these coefficients corresponding to all the resonance lines is equal to $\frac{1}{3}s$ if s is the degree of periodicity of the atom

$$\Sigma f_a = \frac{s}{3}.$$

Digitized by Google

In the foregoing paper the assumption $S = 4p$ has been made, so that

$$\Sigma f_a = \frac{4p}{3}.$$

Only those lines which make an appreciable contribution to the dispersion have been introduced into the formula and the f_a's have been so chosen that for the lines in the formula

$$\Sigma f_a < \frac{4p}{3}.$$

This assumption is the quantum theory analogue of that made by Herr Unsöld.

It should also be pointed out that in considering Aluminium the resonance line $1S - 1p_1$ cannot be ignored as stated on page 96. Although it is a weak line it occurs actually in the range of transition frequencies of the series electron and is therefore important.

In conclusion, I wish to express my thanks to Mr R. H. Fowler and Mr D. R. Hartree for their help and criticism during the course of this work, and to the Board of Scientific and Industrial Research for a grant which enabled me to pursue this course of research.

Digitized by Google

On Quantum Algebra. By Dr P. A. M. DIRAC, St John's College. (Communicated by Mr R. H. FOWLER.)

[*Received* 17 July, *read* 26 July 1926.]

For the purposes of atomic physics it has been found convenient to introduce the idea of quantities that do not in general satisfy the commutative law of multiplication, but satisfy all the other laws of ordinary algebra*. These quantities are called q-numbers, and the numbers of ordinary mathematics c-numbers, while the word number alone is used to denote either a q-number or a c-number. Both q-numbers and c-numbers can occur together in the same piece of analysis, and even in the same equation, as numbers of the two kinds can be added together or multiplied. A c-number may, in fact, be regarded as a special case of the more general q-number. In the particular case when two numbers x and y satisfy $xy = yx$, we shall say that x commutes with y. A c-number is assumed to commute with every number.

The development of the algebra of q-numbers has up to the present been greatly hampered by the fact that there has been no general definition of a function of a q-number variable. One has been obliged to restrict the word function to apply only to analytic functions specified by algebraic equations involving the operations of addition, multiplication and reciprocation (which last applied to q-numbers can easily be given a definite meaning), and one has usually contented oneself with taking a theorem involving arbitrary functions to be true generally when one has proved it true for power series.

In the present paper a general definition of a function is proposed which appears to enable one to establish theorems involving arbitrary functions with rigour. The differential coefficient of a function of a single variable can also be conveniently defined, its most important properties being contained in, or easily deducible from, its definition.

Functions of a q-number variable.

Consider a simple algebraic function of a single q-number variable x, such as x^2. Evidently every number that commutes with x must also commute with x^2. This property provides us with the general definition of a function of a q-number variable.

Let x and X be two q-numbers such that any number that commutes with x also commutes with X. We then say that X is a function of the single variable x, and consider the two numbers

* *Roy. Soc. Proc.* A, vol. 110, p. 561 (1926).

Digitized by Google

X and x as specifying a function. A function of a q-number variable is thus specified by a pair of q-numbers that satisfy a certain condition and is quite different from a function of a real c-number variable, for the specification of which an infinite number of pairs of numbers are required, which need not satisfy any conditions.

In order that we may consider the two q-numbers x and X as really specifying a function, we must be able to say what is the value of this function for any particular value of the variable, y say. We take the rule for this to be that if b satisfies the equation

$$y = bxb^{-1} \quad \text{.....(1),}$$

then Y defined by

$$Y = bXb^{-1} \quad \text{.....(2)}$$

is the same function of y that X is of x. This rule completes the definition of a function of a single q-number variable. (It is taken as an axiom that there always exists a q-number b satisfying (1) when x and y are q-numbers.)

In order that our definition of a function may be tenable, one must be able to deduce the following theorems:

(i) All values of b that satisfy (1) must give the same value for Y defined by (2).

(ii) Y defined by (2) must be a function of the single variable y, *i.e.* any number that commutes with y must also commute with Y.

(iii) Functions of a single variable can be formed by the addition or multiplication of functions of that variable.

(iv) A function of a function of a certain variable must be a function of that variable.

To prove (i) we observe that if b' is another root of (1), we have

$$bxb^{-1} = y = b'xb'^{-1},$$

whence, dividing by b in front and multiplying by b' behind, we get

$$xb^{-1}b' = b^{-1}b'x.$$

Thus $b^{-1}b'$ commutes with x, and therefore also with X, since by hypothesis every number that commutes with x also commutes with X. Hence

$$b^{-1}b'X = Xb^{-1}b',$$

so that, multiplying by b in front and dividing by b' behind, we find

$$b'Xb'^{-1} = bXb^{-1} = Y$$

from (2), which proves the theorem.

Digitized by Google

To prove (ii), let a be any number that commutes with y, so that

$$ay = ya,$$

or

$$abxb^{-1} = bxb^{-1}a,$$

from (1). This gives

$$b^{-1}abx = xb^{-1}ab,$$

so that $b^{-1}ab$ commutes with x, and hence also with X. Thus

$$b^{-1}abX = Xb^{-1}ab,$$

which gives

$$abXb^{-1} = bXb^{-1}a,$$

or

$$aY = Ya,$$

which proves the theorem.

To prove (iii), we observe that if X_1 and X_2 are two functions of the single q-number variable x, then $X_1 + X_2$ and X_1X_2 are also functions of the single variable x, since any number that commutes with x must commute with X_1 and X_2, and therefore commutes with $X_1 + X_2$ and X_1X_2. To complete the proof that functions of a single variable may be formed by the addition or multiplication of functions of that variable, we must show that if Y_1 and Y_2 are the same functions of y that X_1 and X_2 are of x, then $Y_1 + Y_2$ and Y_1Y_2 are the same functions of y that $X_1 + X_2$ and X_1X_2 are of x. By hypothesis we have

$$y = bxb^{-1}, \qquad Y_1 = bX_1b^{-1}, \qquad Y_2 = bX_2b^{-1},$$

and hence

$$Y_1 + Y_2 = b(X_1 + X_2)b^{-1},$$

$$Y_1Y_2 = bX_1b^{-1}.bX_2b^{-1} = bX_1X_2b^{-1},$$

which proves the theorem.

Finally, to prove (iv), we see that if X is a function of the single variable x and ξ is a function of the single variable X, then any number that commutes with x must commute with X, and hence must also commute with ξ, so that ξ is a function of the single variable x. Also, if Y is the same function of y that X is of x, we have

$$y = bxb^{-1}, \qquad Y = bXb^{-1},$$

and if in addition η is the same function of Y that ξ is of X, we have

$$\eta = b\xi b^{-1}.$$

It follows that η is the same function of y that ξ is of x.

Theorem (iii) shows that algebraic functions defined by the operations of addition and multiplication are still functions in the new sense of the word. Also it is easily verified that x^{-1} (defined by $x^{-1}x = 1$ or $xx^{-1} = 1$) is a function of x, and hence from (iv) the reciprocal of any function of x is a function of x. Any function of

Digitized by Google

a single variable must commute with any other function of that variable. A c-number may be regarded as a particular case of a function of any q-number variable.

The extension of the theory to functions of more than one variable is rather restricted. We define X to be a function of the variables $x_1, x_2, \ldots x_n$ only, if X commutes with every number that commutes with $x_1, x_2, \ldots x_n$. If we now have

$$y_1 = bx_1b^{-1},\ y_2 = bx_2b^{-1},\ \ldots\ y_n = bx_nb^{-1} \quad \ldots\ldots\ldots(3),$$

and put

$$Y = bXb^{-1},$$

we can define Y to be the same function of the variables $y_1, y_2, \ldots y_n$ that X is of the variables $x_1, x_2, \ldots x_n$, and can easily deduce the theorems corresponding to (i)...(iv) for functions of several variables. It is, however, only when the y's satisfy the conditions (3) that one can give a meaning to the expression "the same function of the y's that X is of the x's," owing to the fact that one cannot consider the x's to vary independently, as one could if they were c-numbers. Suppose, for instance, that x_1 and x_2 satisfy

$$x_1x_2 - x_2x_1 = 1.$$

Then one would have to count x_1x_2 and $(1 + x_2x_1)$ as two ways of writing the same function of x_1 and x_2; but one would not in general have y_1y_2 equal to $(1 + y_2y_1)$, except when the y's satisfy (3), so that it is only in this case that one can hope to give a meaning to the same function of the y's.

Differential coefficients.

Let Q be a function of the single variable q, and let p be a number that satisfies

$$qp - pq = 1 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4).$$

Put

$$Qp - pQ = Q' \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(5).$$

Then we define Q' to be the differential coefficient of Q with respect to q, and write it dQ/dq.

In order that our definition may be tenable, one must be able to deduce the following theorems:

(i) All values of p that satisfy (4) must give the same value for Q' defined by (5).

(ii) Q' defined by (5) must be a function of the single variable q.

(iii) The ordinary rules for the differential coefficient of a sum and a product must still hold.

(iv) The ordinary rule for the differential coefficient of a function of a function must still hold.

Digitized by Google

To prove (i), we observe that if p' is another root of (4), we have

$$qp - pq = 1 = qp' - p'q,$$

or

$$q(p - p') = (p - p')\, q.$$

Thus $(p - p')$ commutes with q, and therefore also with Q. Hence

$$Q(p - p') = (p - p')\, Q,$$

or

$$Qp' - p'Q = Qp - pQ = Q',$$

which proves the theorem.

To prove (ii), let a be any number that commutes with q. It must therefore commute with Q. We have to prove that it commutes with Q'. From (4), since $(qp - pq)$ is a c-number, we have

$$\begin{aligned} 0 &= a(qp - pq) - (qp - pq)\, a \\ &= aqp - apq - qpa + pqa \\ &= qap - apq - qpa + paq \\ &= q(ap - pa) - (ap - pa)\, q. \end{aligned}$$

Hence $(ap - pa)$ commutes with q, and therefore also with Q, so that

$$\begin{aligned} 0 &= Q(ap - pa) - (ap - pa)\, Q \\ &= Qap - Qpa - apQ + paQ \\ &= aQp - Qpa - apQ + pQa \\ &= a(Qp - pQ) - (Qp - pQ)\, a \\ &= aQ' - Q'a. \end{aligned}$$

Thus a commutes with Q'. To complete the proof of (ii), we must show that if Q_1 is the same function of q_1 that Q_2 is of q_2, then Q_1' is the same function of q_1 that Q_2' is of q_2, *i.e.* if

$$q_2 = bq_1b^{-1}, \qquad Q_2 = bQ_1b^{-1},$$

$$q_1p_1 - p_1q_1 = 1, \qquad q_2p_2 - p_2q_2 = 1,$$

and

$$Q_1p_1 - p_1Q_1 = Q_1', \qquad Q_2p_2 - p_2Q_2 = Q_2',$$

then

$$Q_2' = bQ_1'b^{-1}.$$

We have

$$1 = q_2p_2 - p_2q_2 = bq_1b^{-1}p_2 - p_2bq_1b^{-1},$$

from which, by multiplying by b behind and b^{-1} in front, we obtain

$$1 = q_1b^{-1}p_2b - b^{-1}p_2bq_1.$$

Digitized by Google

This shows that $b^{-1}p_2 b$ may be taken to be p_1. Hence

$$Q_1' = Q_1 b^{-1} p_2 b - b^{-1} p_2 b Q_1,$$

or

$$bQ_1'b^{-1} = bQ_1 b^{-1} p_2 - p_2 bQ_1 b^{-1}$$
$$= Q_2 p_2 - p_2 Q_2 = Q_2',$$

which proves the theorem.

To prove (iii) let Q_1 and Q_2 be two functions of the single variable q. It is obvious that

$$(Q_1 + Q_2)' = Q_1' + Q_2'.$$

Further

$$(Q_1 Q_2)' = Q_1 Q_2 p - pQ_1 Q_2$$
$$= (Q_1 p - pQ_1) Q_2 + Q_1 (Q_2 p - pQ_2)$$
$$= Q_1' Q_2 + Q_1 Q_2'.$$

The ordinary laws for the differentiation of a sum and a product thus still hold.

To prove (iv), let Q be a function of the single variable q and ξ a function of the single variable Q, and let

$$qp - pq = 1,$$

and

$$QP - PQ = 1 \quad \ldots\ldots(6).$$

We then have by definition

$$Qp - pQ = dQ/dq \quad \ldots\ldots(7),$$

$$\xi p - p\xi = d\xi/dq \quad \ldots\ldots(8),$$

and

$$\xi P - P\xi = d\xi/dQ \quad \ldots\ldots(9).$$

Now dQ/dq must commute with Q, since they are both functions of the single variable q, and hence, dividing (7) by dQ/dq in front, we obtain

$$1 = \left(\frac{dQ}{dq}\right)^{-1}(Qp - pQ) = Q\left(\frac{dQ}{dq}\right)^{-1} p - \left(\frac{dQ}{dq}\right)^{-1} pQ.$$

This shows that $(dQ/dq)^{-1}p$ may be taken to be the P of equation (6). Hence, using this value for P in equation (9), we obtain

$$\frac{d\xi}{dQ} = \xi\left(\frac{dQ}{dq}\right)^{-1} p - \left(\frac{dQ}{dq}\right)^{-1} p\xi = \left(\frac{dQ}{dq}\right)^{-1}(\xi p - p\xi) = \left(\frac{dQ}{dq}\right)^{-1}\frac{d\xi}{dq}$$

using (8) and the fact that $(dQ/dq)^{-1}$ commutes with ξ. Hence

$$\frac{d\xi}{dq} = \frac{dQ}{dq}\,\frac{d\xi}{dQ},$$

which is the ordinary rule for the differentiation of a function of a function.

From (iii) and (iv) it follows that the differential coefficients of simple algebraic functions have their ordinary values. In order

Digitized by Google

that the converse process of integration applied to simple algebraic functions may also lead to the usual results, it is necessary and sufficient that the only functions whose differential coefficients are zero shall be constants (*i.e.* c-numbers). This requires that the only numbers that commute with p and also with every number that commutes with q, when p and q satisfy (4), shall be c-numbers. We take this to be an axiom in the new algebra. It is a particular case of a more general axiom which may be enunciated thus: If a number commutes with every number that commutes with the q-numbers q, q_1, q_2, ... q_n and with one number that does not commute with q but commutes with $q_1, q_2, \ldots q_n$, then it commutes with every number that commutes with $q_1, q_2, \ldots q_n$*. The axiom may also be stated geometrically thus: Each q-number may be represented by a point in space of an infinite number of dimensions in such a way that the condition that a q-number shall commute with certain given q-numbers is that it shall lie in a certain flat space.

The definition of a differential coefficient may be extended to the partial differential coefficients of a function of several variables that commute with one another. If Q is a function of the variables q, q_1, q_2, ... q_n that commute with one another, to define $\partial Q/\partial q$ one requires a p that satisfies (4) and that commutes with $q_1, q_2, \ldots q_n$. One can then define $\partial Q/\partial q$ by the left-hand side of (5), and can show that it satisfies theorems corresponding to (i)...(iv). It is impossible, however, to extend the definition to the partial differential coefficients of functions of several arbitrary variables. Suppose, for example, that x_1 and x_2 are q-numbers such that

$$x_1 x_2 = x_2 (x_1 + c),$$

where c is a c-number; then $x_1 x_2$ and $x_2 (x_1 + c)$ must be two ways of writing the same function of x_1 and x_2, and it would obviously be impossible to give a rational definition of the partial differential coefficient of this function with respect to x_2.

* Note that a number x that commutes with every number must be a c-number, on account of the axiom that if x is a q-number, there must be a q-number b that makes bxb^{-1} equal an arbitrary q-number.

Digitized by Google

An optical illusion due to contrast. By Mr G. P. THOMSON, Corpus Christi College.

[*Read* 26 July 1926.]

(Plate III.)

In the course of a long series of experiments on the scattering of positive rays I have had occasion to take some hundreds of photographs in which positive rays diverging from an approximately linear source pass through a metal slit about ·5 mm. wide parallel to the source and then strike the sensitive surface of the plate. The change caused by the actual impact of the rays can be developed like the ordinary latent image produced by light. The pattern thus formed on the negative is a rectangular strip about 1·8 mm. × 20 mm. answering in shape to the shape of the slit, and bordered by a much less intense edge (of the order 10 °/₀) due to the scattered rays whose investigation was the object of the experiment. It was found that, in a majority of the photographs, the rectangular strip, which should have been approximately uniform in blackness, at least across, had its longer edges apparently much blacker than its centre. So marked is this effect that it was supposed to be due to some cause affecting the distribution of the rays. However, careful measurements with a microphotometer showed that the effect was unreal, the edges always being less black than the centre. The effect of the positive rays is thus merely to produce a certain distribution of blackening on the plate, and the increased blackness seen at the edges is purely subjective.

The effect was sufficiently striking to make it seem worth further investigation. It was found that the effect persisted under ordinary photographic enlargement to about four times the linear dimensions. When a positive (enlarged) was made of two of the plates, the converse effect was seen, the edges of the bright strip now appearing brighter than the centre, the strip seeming to have a dark band down its centre. If anything this effect was more marked than in the original negatives. The upper figures in Plate III are positives from these negatives, the lower are negatives from the original positives. The further photographic processes involved in reproduction cause of course a loss of distinctness.

Among the large number of negatives taken the effect was shown in very varying degrees, though there were few in which it was entirely absent. It soon appeared that this was due to the varying extent to which the rectangular pattern was blurred by the scattering radiation which varied considerably in amount from one experiment to another.

Digitized by Google

Where there was little scattering, the enhanced edges were very distinct, but narrow. As the scattering increased the edges appeared broader but less distinct, until the effect appeared only as a slightly lighter centre line to the black band, the existence of which was a matter of opinion, and finally disappeared altogether, leaving a strip visibly darker in the centre.

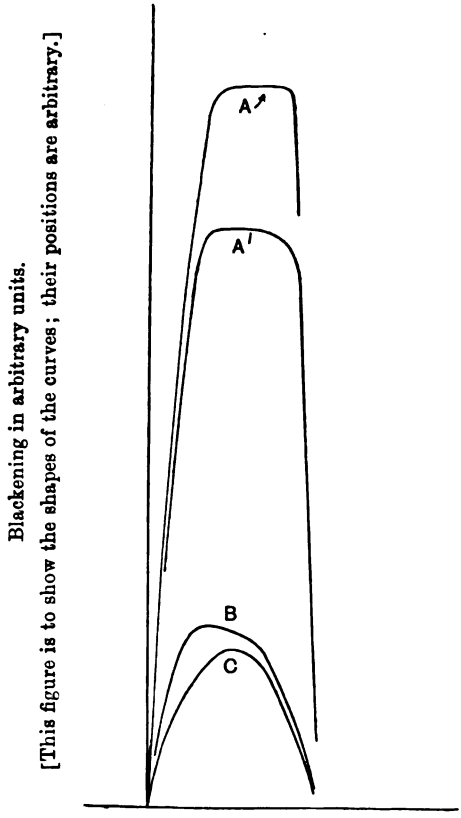


Fig. 1.

The natural interpretation of these results is that the eye sees rapid *change* of blackness*, on a light background, as enhanced blackness. Thus in the case of little scattering, the blackening curve as measured by the photometer (see fig. 1, A, A') is flat-topped. There are two narrow regions near the edge where the

* Or, what would fit these experiments equally well, a place of sharp curvature on the blackening curve.

Digitized by Google

Digitized by Google

Digitized by Google

change of blackening is very rapid, and these appear as narrow edges of enhanced blackening. As the pattern becomes more blurred the slope at the edges becomes more gentle and extends over a wider region (fig. 1, *B*). Thus the enhanced edges grow wider but less distinct, until finally we reach the condition shown in fig. 1, *C*, where the central (approximately) flat portion disappears, and the illusion ceases. Actually in the plate measured for (*C*) under some conditions of light and angle the centre appears blackest, under others it appears as slightly lighter than the edges. The curve in fact represents the critical stage at which the illusion ceases.

To test this hypothesis negatives were taken in which a rectangular slit was placed close to the plate. The edges of the pattern formed were thus quite sharp and the blackening curve of a cross-section of the pattern would be flat-topped with sharp corners. In this case no effect was observed, the strip appearing uniform. On slightly separating the slit from the plate a fine line appeared at the edges, so fine as to require fairly careful inspection. This corresponds to the narrow region of "penumbra" introduced by the separation of plate and slit. These effects occur equally with positive rays and ordinary light.

The existence of the effect in the positives shows that rate of change of brightness has a similar influence on the appearance of a light object on a dark ground. It should be noticed that this is not a case of irradiation in the ordinary sense, as it is not a question of the apparent *size* of an object being modified by contrast, but the relative brightness of its different parts.

Another effect of the same kind which sometimes appears, though much more rarely, is the appearance of a white line between the main portion of the pattern and the much fainter scattered region, although the curve of blackening shows no minimum. Here there is the complication of a region of faint blackening between two opposite and contrasting backgrounds. If the eye considers that part of the faint object which is near the blackened side as a light object on a dark ground, then the appearance of a white line at the place of most rapid change is analogous to the bright edges to the positive in the plate.

Dr C. D. Ellis tells me that he has found effects of both these kinds in the magnetic spectra of β rays. I am much indebted to him and Mr Wooster for their kindness in helping me to use the microphotometer in the Cavendish Laboratory, on which the measurements of fig. 1 were made, and to Sir Ernest Rutherford for permission to use it.

Digitized by Google

On the Quantum Theory of the Problem of the Two Bodies. (Preliminary Communication.) By J. R. OPPENHEIMER. (Communicated by Mr R. H. FOWLER.)

[*Read* 26 July 1926.]

The problem of the two bodies has been treated on the new mechanics by Dirac*, Pauli†, and Schrödinger‡, who have independently derived the Balmer terms. The present paper is an attempt at a more complete solution. In particular, formulae are derived for the line intensities of the hydrogen spectrum, for the photoelectric effect and its inverse, and for the continuous absorption spectrum in the ultraviolet and in the X-ray regions. Also the probabilities of transition, deflection and capture are computed for the collision of an electron and an ion. Numerical values are only obtained, however, for the simplest line intensities. It is hoped to treat the problem in greater detail.

I. *Elliptic Orbits.* For the line intensities the system may be treated as multiply periodic. We shall assume three degrees of freedom, and characterize transitions by six quantum numbers $(nkm, n'k'm')$. The fourth degree of freedom which would follow from the existence of an electronic magnetic moment, would not affect the intensities of the series lines as wholes. For the emission lines we may use the classical formula§ for the intensity in which frequencies and amplitudes have been replaced by their quantum theoretic analogues‖. Thus

$$I(nkm, n'k'm') = \sum_{\alpha=1}^{3} \frac{2^4\pi^4\epsilon^2}{3c^3} [\nu^4(n, n')] |x_\alpha(nkm, n'k'm')|^2 \ldots (1).$$

The frequencies are

$$\nu(nn') = R\left(\frac{1}{n^2} - \frac{1}{n'^2}\right) \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2).$$

According to Schrödinger¶ the amplitudes are given by

$$x_\alpha u(nkm) = \sum_{n'} \sum_{k'} \sum_{m'} x_\alpha(nkm, n'k'm')\, u(n'k'm') \ldots\ldots (3),$$

where the functions $u(nkm)$ satisfy the following conditions:

1. They are bounded everywhere within the configuration space $0 \leqslant r$; $0 \leqslant \vartheta \leqslant \pi$; $0 \leqslant \phi \leqslant 2\pi$, and tend to zero sufficiently fast as $r \to \infty$.

* Dirac, *Proc. Roy. Soc.* A, 110, 561 (1926).
† Pauli, *Zeit. f. Phys.* 36, 336 (1926).
‡ Schrödinger, *Ann. d. Phys.* 79, 361 (1926).
§ Kramers, *Intensities of Spectral Lines*, Copenhagen, 1919.
‖ Born u. Jordan *Zeit. f. Phys.* 34, 858 (1925).
¶ Schrödinger, *Ann. d. Phys.* 79, 734 (1926).

Digitized by Google

2. They are twice continuously differentiable.

3. They satisfy, if μ, ϵ and E are mass, charge and energy of the electron

$$\nabla^2 u + \frac{8\pi^2\mu}{h^2}\left(E + \frac{\epsilon^2}{r}\right)u = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots (4).$$

4. They are one-valued within the configuration space.

5. They are normal and orthogonal with respect to the configuration space

$$\iiint u(nkm)\, u(n'k'm')\, \rho\, dr\, d\vartheta\, d\phi = \delta_n^{n'}\, \delta_k^{k'}\, \delta_m^{m'}\ ;\ \rho = \frac{\partial(x_1 x_2 x_3)}{\partial(r\phi\vartheta)} = r^2 \sin\vartheta \quad \ldots\ldots\ldots\ldots (5).$$

From this it follows that

$$x_a(nkm,\, n'k'm') = \int_0^\infty \int_0^\pi \int_0^{2\pi} u(nkm)\, u(n'k'm')\, x_a(r\vartheta\phi)\, \rho(r\vartheta\phi)\, d\phi\, d\vartheta\, dr \ldots (6).$$

For $E < 0$, it is only possible to construct solutions which have the form

$$u(nkm) = u_{nkm}(r, \vartheta, \phi) = A \begin{matrix}\cos m\phi \\ \sin m\phi\end{matrix} P_{\bar{k}}^{(m)}(\cos\vartheta)\, \alpha^{\bar{k}} e^{-\alpha} \sum_{j=0}^{n-1-\bar{k}} \frac{(-2\alpha)^j}{j!} \binom{n+\bar{k}}{n-\bar{k}-1-j} \quad \ldots\ldots\ldots\ldots (7),$$

where $\bar{k} = k - \frac{1}{2}$ and $\alpha = (4\pi^2\mu\epsilon^2/nh^2)r$ satisfy (1)—(5) for certain values of the constants, namely

$$\left.\begin{array}{l} n = 1,\quad 2,\quad 3 \ldots\ldots\ , \\ \bar{k} = 0,\quad 1,\quad 2 \ldots\ldots (n-1) \\ m = 0,\ \pm 1,\ \pm 2 \ldots\ldots \pm \bar{k} \end{array}\right\} \ldots\ldots\ldots\ldots\ldots (8).$$

From (1), (6) and (7), we get, on performing the integrations, the selection rules

$$\left.\begin{array}{l} m - m' = 0,\ \pm 1 \\ k - k' = \pm 1 \end{array}\right\} \ldots\ldots\ldots\ldots\ldots\ldots\ldots (9).$$

Averaging over m and m', we get the component intensities*

$$I(nk, n'k') = \frac{2^7\pi^6\mu^3\epsilon^{14}}{3c^3h^8k}\left(\frac{1}{n'^2} - \frac{1}{n^2}\right)^4 \frac{(k+k')\, A^2(nk, n'k')}{A^2(nk)\, A^2(n'k')} \ldots (10).$$

* The intensities of the hydrogen lines have been investigated by Pauli in a paper not yet published.

Digitized by Google

where

$$A^2(nk) = \sum_{\gamma_1=0}^{n-\bar{k}-1} \sum_{\gamma_2=0}^{n-\bar{k}-1} \frac{n^3(-1)^{\gamma_1+\gamma_2}(2k+\gamma_1+\gamma_2+1)!\,[(n+\bar{k})!]^2}{\gamma_1!\,\gamma_2!\,(2k+\gamma_1)!\,(2k+\gamma_2)!\,2^{2k}(n-\bar{k}-1-\gamma_1)!\,(n-\bar{k}-1-\gamma_2)} \quad \ldots\ldots\ldots\ldots(11)$$

and

$$A(nk, n'k') = \sum_{\gamma_1=0}^{n-\bar{k}-1} \sum_{\gamma_2=0}^{n'-\bar{k}'-1} \frac{(nn')^4 n^{\bar{k}+\gamma_1} n^{\bar{k}+\gamma_2}(-2)^{\gamma_1+\gamma_2}(n+\bar{k})!\,(n'+\bar{k}')!\,(k+k'+\gamma_1+\gamma_2+}{(n+n')^{k+k'+\gamma_1+\gamma_2+3}\gamma_1!\,\gamma_2!\,(2k+\gamma_1)!\,(2k'+\gamma_2)!\,(n-\bar{k}-1-\gamma_1)!\,(n'-\bar{k}'-1-\gamma_2)!} \quad \ldots\ldots\ldots\ldots(12)$$

Averaging again over k and k' we get the series intensities. In ergs per atom per second these give*, for the first and second Lyman lines, 0·00667 and 0·000644; for the first and second Balmer lines, 0·00789 and 0·00199. The only directly observable ratio is that of the second Lyman to first Balmer line. Thus

$$\frac{I(3, 2)}{I(3, 1)} = 12{\cdot}2.$$

The absorption lines (n, n') from the initial state n, have total intensities†

$$\frac{n'^2}{n^2}\frac{c^3h^8}{2\pi^2\epsilon'^2\mu^3}\left(\frac{1}{n^2}-\frac{1}{n'^2}\right)^{-3} \quad \ldots\ldots\ldots\ldots(13)$$

times the corresponding emission intensities. This gives, for instance, the two observable ratios

$$\frac{I_{abs}(1, 2)}{I_{abs}(1, 3)} = 7{\cdot}72; \quad \frac{I_{abs}(2, 3)}{I_{abs}(2, 4)} = 5{\cdot}37.$$

The evidence in favour of such calculations of absolute intensities is not very extensive. Two such calculations were carried through on the Correspondence Principle. Thus Tolman‡ found an approximate agreement for the outer lines of a vibration-rotation band in the infra red; and Kramers§ found a rough value for the efficiency of an X-ray bulb.

II. *Hyperbolic Orbits.* It is possible to extend this method to the treatment of hyperbolic orbits. Classically the Fourier coefficients for the coordinates do not exist for these orbits. On

* I am indebted to Mr J. T. Edsahl for checking these calculations.
† Fowler, *Phil. Mag.* 50, 1079 (1925).
‡ Tolman and Badger, *Phys. Rev.* 27, 383 (1926).
§ Kramers, *Phil. Mag.* 46, 836 (1923).

Digitized by Google

the quantum theory the components corresponding to a transition between two hyperbolic orbits do not exist and we are forced to use the accelerations directly. For photoelectric transitions, however, the coordinate components may be found, and the intensities are derived from them in accordance with (1). Here, however,

$$\nu(n, E) = \frac{R}{n^2} + \frac{E}{h} \quad \text{.....(14),}$$

where E is the positive energy of the hyperbolic orbit. The factors in the solution of (4) which depend on ϑ and ϕ are those given in (7). The values of k and m for which these satisfy the conditions (1)—(5), are

$$\left.\begin{aligned} k &= \tfrac{1}{2}, \quad \tfrac{3}{2}, \quad \tfrac{5}{2} \ \ldots\ldots \ \ldots\ldots \\ m &= 0, \pm 1, \pm 2 \ \ldots\ldots \pm \bar{k} \end{aligned}\right\} \quad \text{.....(15).}$$

For all values of m and k it is possible to find an integral function $f(E, k)$ of r satisfying (4), and which, for a set of values of E which become closer and closer as the configuration space is expanded, satisfy the conditions (1) to (4), and are orthogonal. The characteristic functions so obtained are

$$u(Ekm) = A_1 \begin{matrix}\cos m\phi \\ \sin m\phi\end{matrix} P_{\bar{k}}^{(m)}(\cos\vartheta) f(Ek) \quad \text{.....(16),}$$

where

$$f(Ek) = f_{Ek}(r) = \sum_{\nu=0}^{\infty} c_\nu(Ek)\, r^{\nu+\bar{k}} \quad \text{.....(17).}$$

The coefficients are given, on the one hand, by the double circuit integrals†

$$2c_\nu(Ek) = \left[\int_{(\gamma,-\gamma)} (z+\gamma)^{\alpha_1-1}(z-\gamma)^{\alpha_1-1}\, dz\right]^{-1} \times B(\alpha_1\alpha_2) \int_{(\gamma,-\gamma)} \frac{z^\nu}{\nu!}(z+\gamma)^{\alpha_2-1}(z-\gamma)^{\alpha_1-1}\, dz \quad \text{.....(18),}$$

where

$$\left.\begin{aligned} \alpha_1 &= \bar{k}+1-in^*; \quad \alpha_2 = \bar{k}+1+in^* \\ n^* &= (2\pi^2\mu\epsilon^4/h^2E)^{\frac{1}{2}}; \ \gamma = \frac{2\pi i}{h}(2\mu E)^{\frac{1}{2}} \end{aligned}\right\} \quad \text{.....(19)}$$

and on the other, by the recurrence formulae

$$\left.\begin{aligned} &(\nu^2+2k\nu)\, c_\nu + 8\pi^2 m h^{-2}(\epsilon^2 c_{\nu-1} + E c_{\nu-2}) = 0 \\ &c_0 = \frac{1}{2} B(\alpha_1, \alpha_2) = \frac{\Gamma(\alpha_1)\,\Gamma(\alpha_2)}{\Gamma(\alpha_1+\alpha_2)} \end{aligned}\right\} \quad \text{...(20).}$$

† See, for an account of the mathematical methods, Schlesinger, *Differentialgleichungen*, Vieweg, 1922, and Whittaker and Watson, *Modern Analysis*, Cambridge, Chap. XIV.

Digitized by Google

From these it follows that

$$c_\nu(Ek) = S_\nu(E)\left\{\prod_{\tau=1}^{\nu}(\tau^2+2k\tau)\right\}^{-1} \quad \text{.....(21),}$$

where S_ν is of degree $\left[\frac{\nu}{2}\right]$ in E, and thus that the series defining f is absolutely convergent and real. Finally, f is given asymptotically for large values of $(2\gamma r)$ by the expansion

$$f(Ek) \sim r^{\bar{k}}(-2\gamma r)^{-\alpha_1} e^{\gamma r}\Gamma(\alpha_1)\sum_{s=0}^{\infty}\frac{\Gamma(\alpha_1+s)}{\Gamma(s)\,\Gamma(\alpha_1-s)}(-2\gamma r)^{-s}$$

$$-r^{\bar{k}}(-2\gamma r)^{-\alpha_2} e^{-\gamma r}\Gamma(\alpha_2)\sum_{s=0}^{\infty}\frac{\Gamma(\alpha_2+s)}{\Gamma(s)\,\Gamma(\alpha_2-s)}(+2\gamma r)^{-s} \quad \text{...(22),}$$

which only fails when E is zero.

It is now possible to see in detail how this case differs from the periodic one. In the first place it follows from (22) that we can no longer choose the constant s_1 to make

$$\int_0^\infty\int_0^\pi\int_0^{2\pi} u^2\rho\,d\phi\,d\vartheta\,d\nu = 1.$$

The functions u are thus still orthogonal, but cannot be made normal. In the second place, even if we eliminate all values of E for which condition 1 is not satisfied, we are left with a continuous manifold of characteristic functions. The equations (3) thus become meaningless, and must be replaced by

$$\left.\begin{aligned} u(nkm)\,x_\alpha &= \sum_{n'}\sum_{k'}\sum_{m'} x_\alpha(nkm, n'k'm')\,u(n'k'm') \\ &\quad + \int_0^\infty \sum_{k'}\sum_{m'} x_\alpha(nkm, Ek'm')\,df(E)\,.\,u(Ek'm') \\ u(Ekm)\,x_\alpha &= \sum_{n'}\sum_{k'}\sum_{m'} x_\alpha(Ekm, n'k'm')\,u(n'k'm') \\ &\quad + \int_0^\infty \sum_{k'}\sum_{m} x_\alpha'(Ekm, E'k'm')\,u(E'k'm')\,.\,df(E') \end{aligned}\right\} \quad \text{.....(23).}$$

Here the u_n are to be normalized, and the u_E chosen to make

$$x_\alpha(Ekm, nk'm') = x_\alpha'(nk'm', Ekm) \text{.....(24).}$$

Finally we may determine $f(E)$ by the following argument.

Classically we should have, for the acceleration of the electron in an orbit of energy E,

$$\ddot{x}_\alpha(E) = \int_{-\infty}^{\infty}\ddot{x}_\alpha(E,\nu)\exp(2\pi i\nu t)\,d\nu,$$

and, for the energy radiated from that orbit

$$\text{const.}\sum_{\alpha=1}^{3}\int_{-\infty}^{\infty}\ddot{x}_\alpha^2(E\nu)\exp(2\pi i\nu t)\,d\nu.$$

Digitized by Google

Quantum theoretically we have, for the components of acceleration corresponding to a photoelectric transition

$$\left.\begin{aligned} u(n)\ddot{x}_\alpha &= \int_0^\infty 4\pi^2\nu^2(nE)\,x_\alpha(nE)\,u(E)\exp[2\pi i(E/h)t]df(E) \\ &= \int_0^\infty \left[4\pi^2\nu^2(nE)\,x_\alpha(nE)\frac{\partial f(E)}{\partial E}.h\right]u(E)\exp[2\pi i(E/h)t]d(E/h) \end{aligned}\right\} \quad \text{.....(25)}$$

[we suppress the indices k and m throughout this argument].

By analogy the radiated energy should be

$$\text{const.}\sum_\alpha \int_0^\infty \left[hx_\alpha(nE)\,\nu^2(nE)\frac{\partial f(E)}{\partial E}\right]^2 u(E)\exp[2\pi i(E/h)t]d(E/h),$$

so that the energy radiated per unit time from the transition $n \rightarrow (E,\ E+dE)$ is, for given $k+m,\ k'+m'$,

$$\text{const.}\sum_\alpha \left[hx_\alpha(nE)\,\nu^2(nE)\frac{\partial f}{\partial E}\right]^2 d(E/h) \text{ .....(26).}$$

On the other hand the total radiation from inverse photoelectric transitions from an orbit of energy E is

$$\text{const.}\sum_\alpha \sum_n x_\alpha^2(En)\,\nu^4(En).$$

Of this

$$\text{const.}\sum_\alpha x_\alpha^2(En)\,\nu^4(En) \text{ .....(27)}$$

is to be ascribed to transitions $E \rightarrow n$.

Now for the fully excited atom, where all temperature factors are ironed out, the number of atoms in any range of phases will be proportional to the corresponding volume of phase space. Thus if the number of atoms in an elliptic orbit with quantum numbers nkm is one, the number with a given m and k and positive energy between E and $E+\Delta E$ is

$$\frac{1}{h}\int_{\text{cell}} \Delta J\,dw = \int_{\text{cell}} \Delta(E/h)\frac{\partial J}{\partial E}dw \text{ .....(28).}$$

Here J is an action variable for the system and w the canonically conjugate angle variable. The rate at which this number is increasing is, since J is a constant,

$$w\frac{\partial J}{\partial E}\Delta(E/h) \text{ .....(29).}$$

It is easy to shew that

$$w = \frac{\partial E}{\partial J} \text{ .....(30).}$$

Digitized by Google

Thus the number of electrons entering the range of hyperbolic orbits per unit time is

$$\Delta(E/h).$$

This gives, for the condition for detailed equilibrium

$$x_a^2(nE)\,\nu^4(nE)\,\Delta(E/h) = [f'h\,x_a(nE)\,\nu^2(nE)]^2\,\Delta(E/h) \quad \ldots(31),$$

whence

$$f' = \frac{\partial f}{\partial E} = \frac{1}{h}; \quad df = d(E/h) \quad \ldots\ldots\ldots\ldots\ldots(32).$$

We may verify* this result in part by computing how many electrons enter the range of hyperbolic orbits

$$E,\ E+\Delta E;\ k,\ k+\Delta k,\ m,\ m+\Delta m$$

from a free electron gas with Maxwellian distribution-in-velocity. We get

$$\text{const.}\ \Delta E\,.\,\Delta m\,.\,(1-m^2/k^2)^{-\frac{1}{2}}\,k\Delta k\,(k^2+n^2)^{-\frac{1}{2}} \quad \ldots\ldots(33).$$

In the limit of large (E and k) this becomes

$$\text{const.}\ \Delta E\Delta m\Delta k\,(1-m^2/k^2).$$

This agrees with the quantum theoretic calculation except for the factor $(1-m^2/k^2)^{-\frac{1}{2}}$. The experiment of Stern and Gerlach suggests that, in this paradox, the quantum theoretic answer is right.

In any case (32) gives

$$\left.\begin{aligned}(a)\quad u(n)\,x_a &= \sum_{n'=1}^{\infty} x_a(nn')\,u(n') + \int_0^{\infty} x_a(nE)\,u(E)\,d(E+h)\\ (b)\quad u(E)\,x_a &= \sum_{n'=1}^{\infty} x_a(En')\,u(n') + \int_0^{\infty} x_a(EE')\,u(E')\,d(E'+h)\end{aligned}\right\} \quad \ldots\ldots(34)$$

to determine $u(E)$ and thus the $x_a(uE)$. Let us call the factor in u which remains after division by the normalized functions of $\mathfrak{H}$ and ϕ, $\bar{u}$; and let us write

$$\bar{u} = \bar{u}(E, k;\ r) = f(E, k;\ r)\,\chi(Ek) \quad \ldots\ldots\ldots(35).$$

From (3) and (6) we get

$$r(En) = \int_0^{\infty} r^2\,.\,r\,.\,u(E)\,.\,u(n)\,dr = r_0(En)\,\chi(Ek) \quad \ldots(36).$$

Also,

$$r(nn') = \int_0^{\infty} u(n)\,u(n')\,r^2\,.\,r\,.\,dr \quad \ldots\ldots\ldots\ldots(37),$$

where, for (36) and (37)

$$\int_0^{\infty} r^2 u^2(n)\,dr = 1.$$

* I am much indebted to Dr P. A. M. Dirac for suggesting this procedure.

Digitized by Google

Then we get, from (34), (36), (24) and (37), the integral equation for χ

$$u(n)\,r - \sum_{n'=1}^{\infty} r(nn')\,u(n') = \int_0^{\infty} \chi^2(Ek) f(Ek;\, r)\, r_0(En)\, d(E/h) \quad \ldots\ldots(38),$$

which is of the form

$$\int_0^{\infty} \chi^2(En)\, F_1(r;\, E, n, k, k')\, d(E/h) = F_2(r;\, n, k') \ldots(39).$$

The existence of a solution (38) independent of r is made plausible by the orthogonality of the $u(E)$; the mathematical problem will be treated in detail shortly.

With this notation the continuous absorption spectrum is given by

$$I(nk, nk')\, dE = \frac{2^7 \pi^5 \epsilon^2 \mu\,(k'+k)}{3c^3 h^2 k} \left(\frac{E}{h} + \frac{R}{n^2}\right) \frac{\chi^2(Ek)\,\alpha^2}{\beta}\, dE \quad \ldots(40),$$

where

$$\alpha = \sum_{\nu=0}^{\infty} \sum_{\gamma=0}^{n-\bar{k}-1} \left(-\frac{4\pi^2\mu\epsilon^2}{nh^2}\right)^{-(\nu+k+k'+j+3)} c_\nu(Ek') \frac{(-2)^j}{\gamma!} \binom{n+\bar{k}}{n-\bar{k}-1-j} (\nu+k+k'+j+2)!$$

$$\beta = \left(\frac{4\pi^2\mu\epsilon^2}{h^2}\right)^{-2\bar{k}-3} A^2(n, k)$$

and $$|k - k'| = 1.$$

Moreover, if $\rho_\nu(\nu)$ is the energy density of radiation, the distribution of photoelectrons in velocity per atom per unit of time is

$$N(v)\, dv = \frac{2^7 \pi^5 \epsilon^2 \mu^2}{3c^3 h^3} \sum_{n=0}^{\infty} \sum_{\bar{k}=0}^{\bar{k}=\infty} \sum_{\bar{k}'=0}^{n-1(|k-k'|=1)} (k+k') \frac{\chi^2 \rho \alpha^2}{k'\beta} v\, dv \quad \ldots\ldots(41).$$

Similarly the emission spectrum from hyperbolic to elliptic orbits is, in energy per electron

$$I(Ek, nk') = \frac{2^4 \pi^4 \epsilon^2}{3c^3} \left(\frac{E}{h} + \frac{R}{n^2}\right)^4 \frac{k+k'}{k} \frac{(\alpha\chi)^2}{\beta} \quad \ldots\ldots(42),$$
$$|k - k'| = 1.$$

It follows from (22) that the integrals for the components of the coordinates corresponding to transitions between two hyperbolic orbits diverge. If we compute the components of the acceleration directly from the equations of motion we get

$$\left.\begin{aligned} \ddot{x}_a &= \varpi_a \epsilon^2 \mu^{-1} r^{-2} \\ \varpi_1 = \cos\phi \sin\vartheta;\ \varpi_2 &= \sin\phi \sin\vartheta;\ \varpi_3 = \cos\vartheta \end{aligned}\right\} \ldots\ldots(43),$$

Digitized by Google

from which

$$\epsilon^{-2}\mu\ddot{\varpi}_a(Ekm E'k'm')\,d(E'/h) = d(E'/h)\iiint \sin\Im\varpi_a\, u(Ekm)\, u(E'k'm')\, d\phi\, d\Im\, d\nu \quad \ldots(44).$$

Now

$$\int_0^\infty f(Ek) f(E'k')\, d\nu \backsimeq \sum_{\nu=0}^{\infty} g_\nu R_1^{\nu+k+k'+1};\quad g_\nu = \sum_{\tau=0}^{\nu} \frac{c_\tau(Ek)_\nu\, c_{\nu-\tau}(E'k')}{\nu+k+k'} \quad \ldots\ldots(45),$$

where we must choose R_1 to make

$$2\gamma(E) R_1 >> 1;\quad 2\gamma(E') R_1 >> 1.$$

For small values of E, of the order of a volt, the series converges rapidly.

This gives, for the intensity of the continuous emission spectrum

$$I(Ek, E'k')\, d(E'/h) = \frac{2^4\pi^4\epsilon^2}{3c^3}\,\frac{k+k'}{k}\,\chi^{-1}(Ek)\,\chi^{-1}(E'k')\sum_{\nu=0}^{\infty} g_\nu R_1^{\nu+k+k'+1}\, d(E'/h) \quad \ldots\ldots(46).$$

III. *Kinematics.* From (42) and (46) we can get the probability that an electron, entering an hyperbolic orbit, shall be bound to the ion in an elliptic orbit, or shall emerge in another hyperbolic orbit. These probabilities are equal to the corresponding intensity of radiation divided by the energy radiated for a single transition. The total probability of capture for an electron of energy E and total angular momentum k, averaged over the possible values of m, is thus

$$p_c = \sum_{n'=1}^{n'=\infty}\ \sum_{k'=0}^{n'-1(|k-k'|=1)} \frac{I(Ek, n'k')}{E + Rhn'^{-2}} \quad \ldots\ldots\ldots\ldots(47).$$

Classically an angular momentum k would correspond to an hyperbola with asymptote a distance Δ from the origin

$$\Delta = \frac{\epsilon^2}{2E}\,\frac{k/n^*}{(1+k^2/n^{*2})^{\frac{1}{2}}} \quad \ldots\ldots\ldots\ldots\ldots\ldots(48).$$

Thus an electron of energy E must "fall within a circle of area $\frac{\pi\epsilon^4}{4E^2}$ about the ion." Its probability of capture is

$$P_c(E) = \lim_{K\to\infty} \sum_{k=0}^{K} k/c(Ek).\left\{\sum_{k=0}^{K} k\right\}^{-1} \quad \ldots\ldots\ldots(49).$$

The quantity called, classically, the effective cross-section, is, for such electrons

$$\frac{\pi\epsilon^4}{4E^2} P_c(E) \quad \ldots\ldots\ldots\ldots\ldots\ldots(50).$$

Digitized by Google

Now classically the electron approximates more and more closely to a free electron as r is indefinitely increased. This is not, however, true on the new theory. For it is easy to shew, by setting $e = 0$ in (4), that the characteristics for the free electron are

$$\begin{matrix}\cos\\ \sin\end{matrix} m\phi P_{\bar{k}}^{(m)} \cos\vartheta\,(-i\gamma r)^{-\frac{1}{2}} J_k(-i\gamma r) \quad \ldots\ldots\ldots(51),$$

which, for large r, approach

$$\begin{matrix}\cos\\ \sin\end{matrix} m\phi P_{\bar{k}}^{(m)} (\cos\vartheta)\, R\,[i^{\bar{k}+1} r^{-1} e^{\gamma r}] \quad \ldots\ldots\ldots\ldots(52).$$

On the other hand (22) gives, for large r, characteristics which differ from (52) by having $(\gamma r + in' - h\gamma r)$ in place of (γr).

Now it will in general be possible to expand the self adjoint characteristics $u\sqrt{\rho}$ for large r in terms of the characteristics for electrons moving with velocity $\left(\frac{h\gamma}{2\pi\mu}\right)$ in three normal directions in space. If it should be possible to expand* it in terms of the characteristics of only such electrons as have the same energy, this would solve the problem of the deflection spectrum. For the initial beam we should have such a sum of characteristics as would give an electron moving, say, along the x axis with the initial velocity. During the encounter the statistical distribution of the characteristics will be changed, in accordance with (46). By analysis of the resulting characteristics in terms of the characteristics of free electrons with the corresponding final energy moving along the three normal directions, we can find the relative probability of any deflection and change in energy by the encounter†. This problem will be considered in detail.

I should like to thank Mr R. H. Fowler, F.R.S. and Dr P. A. M. Dirac, for many valuable suggestions.

* This is impossible if one neglects transitions.

† In this connection see a paper by Born, *Zeit. f. Phys.* 37, 863 (1926).

Digitized by Google

On the stability of copper catalysts produced by thermal decomposition. By F. H. CONSTABLE, Ph.D., Fellow of St John's College.

[*Received* 6 August, *read* 25 October 1926.]

The view which the author has expressed regarding catalytic activity as due to frozen groups of atoms with strong specific external fields* is further tested by a study of the stability of preparations of copper at the instant of formation, and during the subsequent life of the catalyst.

The copper salts of the monobasic fatty acids were heated in a current of alcohol vapour, and the birth of the activity and its magnitude were measured by the rate of evolution of hydrogen gas from the reaction vessel. The apparatus by means of which these measurements were made has been previously described†.

Table I (A) shows the result of visual observations on the temperature of decomposition of these salts. Table I (B) gives the results obtained by gradually heating a copper salt, which was finely distributed over a porcelain rod, in a stream of alcohol vapour. As the temperature of decomposition of the salt was slowly passed, the metallic, catalytically active copper was produced, and made its presence known by the hydrogen evolved. The time, temperature and rate of evolution of hydrogen were tabulated (cf. Table I (B)). The temperature at which the catalysis started agrees well with the temperature at which metallic copper becomes visible.

It can be seen that the catalyst obtained is quite stable and the activity attained after six minutes persists for hours. The overlapping of heating and cooling measurements also confirms this evidence of stability. Sometimes the catalyst showed a slowly falling activity, but this was not often met with. It was suggested that in these cases complete removal of the fatty acid from the reaction tube had not occurred, and the subsequent adsorption by the copper poisoned the catalyst. The measurements of the reaction velocity obtained on cooling and reheating are given separately. Since the surface is gradually being covered up the cooling curve shows too great, the heating curve too small a temperature coefficient, the mean value lying near the true value. From such measurements it is easy to see the magnitude of the error caused by deducing the temperature coefficient from the heating or cooling curve alone (Table II (1) and (2)).

* *Proc. Roy. Soc.* A, vol. 110, p. 283 (1926).
† *Proc. Camb. Phil. Soc.* vol. 22, p. 738 (1925).

Digitized by Google

TABLE I (A). *The Thermal Decomposition of Copper Salts.*

Salt	Temperature °C.	Nature of products
Formate	100	Gives off water of crystallisation, and becomes deeper blue
	180	Slight darkening
	190	Darkening marked, and decrepitation of individual crystals
	200	Metallic copper observed
	210	Decomposition complete
Acetate	150	Darkens
	245	Gives off acetic acid in considerable quantity
	260	Traces of copper noticed, if temperature be maintained copper is deposited on the glass containing tube
	270	Decomposition complete
Propionate	275	„ „
Valerate	285	„ „
Oxalate	300	Darkens
	305	Brown copper visible
	315	Complete
Malonate	280	Darkens
	300	Copper visible
	310	Complete
Succinate	270	Tinge of brown
	330	General darkening
	340	Fumes
	345	Rapid change to copper colour
	355	Copper throughout

TABLE I (B). *The Birth of Catalytic Activity by Thermal Decomposition.*

(1) Formate. Activity too small at 190° C. to enable the birth to be detected by the method used.

(2) Acetate.

Time hr. min.	Temperature °C.	Reaction velocity c.c. per min.
11.31	258	0·0
11.33	262	2·3
11.35	267	5·0
11.37	270	7·84
11.39	272	8·40
11.41	272	8·63
(During the hour interval the reaction proceeded at 236° C., the temperature was then raised again.)		
12.56	272	8·50
12.58	272	8·55

Digitized by Google

TABLE I (B) (*cont.*)

Time hr. min.	Temperature °C.	Reaction velocity c.c. per min.
(3) Propionate.		
10.37	261	0·33
10.39	265	0·32
10.41	270	0·40
10.43	272	1·41
10.45	274	2·19
10.47	274	2·78
10.49	275	3·10
10.51	275	3·13
10.53	275	3·14
10.55	275	3·15
10.57	275	3·04
10.59	275	3·16
11.1	275	3·10
11.3	275	3·00
11.5	275	3·10
11.7	275	3·07
11.9	275	2·95
(4) Butyrate.		
10.00	272	0·0
10.2	274	0·0
10.4	275	3·6
10.6	278	4·6
10.8	280	8·4
10.10	282	8·6
10.20	280	8·3
(5) Valerate.		
10.15	270	0·0
10.17	275	0·3
10.19	284	1·5
10.21	286	4·6
10.23	288	5·0

When overlapping occurred the reaction velocity temperature curves were drawn, and in addition the catalyst was activated by alternate oxidation and reduction at 300° C., two additional curves being given in each case (Table III). It was found in the case of copper from the salts of the monobasic fatty acids that the maximum activity was practically reached at the first oxidation and reduction, the activity after the next oxidation and reduction being only very slightly greater. Since the limiting value of the catalytic activity depends only on the mass of copper used as catalyst and not on its origin, the comparison of the initial activity with the limiting activity on activation gives a measure of the surface activity of the original preparation. Such a comparison is more advantageous than a comparison per unit mass of copper since all errors due to varying conditions of experiment with various preparations are eliminated. The results obtained are shown in Table III.

Digitized by Google

TABLE II. *Unstable Preparations.*

(1) From formate.

Temperature °C.	Reaction velocity c.c.s H_2 per min. heating obs.	heating calc.	cooling obs.	cooling calc.
230	0·84	0·88	0·63	0·71
235	1·08	1·07	0·82	0·85
240	1·32	1·29	1·04	1·05
245	1·56	1·55	1·30	1·27
250	1·90	1·86	1·61	1·55
255	2·20	2·19	1·92	1·87
260	2·64	2·62	2·28	2·24
265	3·08	3·08	2·71	2·69
Intercept on logarithmic curve*	0·841	—	0·904	
Mean	—	0·872		
Heat of activation	—	19,800 cals.		

(2) From acetate.

Temperature °C.	Reaction velocity cooling obs.	cooling calc.	heating obs.	heating calc.
220	1·04	1·00	0·74	0·81
225	1·18	1·26	0·94	1·00
230	1·65	1·51	1·21	1·21
235	1·98	1·88	1·53	1·45
240	2·18	2·25	1·84	1·76
245	2·82	2·75	2·26	2·14
250	3·38	3·38	2·68	2·51
255	3·89	3·98	3·02	3·02
260	4·74	4·79	3·56	3·56
265	5·70	5·75	4·11	4·25
270	7·08	6·93	4·74	5·01
Intercept on logarithmic curve	0·89	—	0·84	
Mean	—	0·865		
Heat of activation	—	19,700 cals.		

* Intercept on the $\log_{10} v$ axis for the interval $\left\{\frac{1}{T_1} - \frac{1}{T_2}\right\} = 0{\cdot}00020$ obtained from the straight line, $\log_{10} v$ plotted against $1/T$.

Digitized by Google

TABLE III. *Stable Preparations.*

(1) From Formate.

No. of alternate oxidations and reductions	0		1		
	Reaction velocity				
Temperature °C.	obs.	calc.	obs.	calc.	v_0/v_1 ratio
230	0·42	0·41	0·68	0·68	0·62
235	0·50	0·50	0·81	0·83	0·62
240	0·60	0·62	1·00	1·01	0·66
245	0·74	0·73	1·24	1·23	0·60
250	0·88	0·89	1·50	1·47	0·61
255	1·06	1·06	1·78	1·77	0·60
260	1·27	1·32	2·08	2·09	0·60
265	1·58	1·53	2·43	2·46	0·60
Relative initial activity					mean 0·61
Intercept on logarithmic curve		0·89		0·87	
Heat of activation		20,400		19,000 cals.	

(2) From Acetate.

No. of alternate oxidations and reductions	0	1	
	Reaction velocity*		
Temperature °C.	obs.	calc.	v_0/v_1 ratio
220	—	2·05	—
225	—	2·46	—
230	—	3·00	—
235	2·36	3·54	0·67
240	2·78	4·16	0·67
245	3·28	4·88	0·67
250	3·88	5·84	0·67
255	4·46	6·84	0·65
260	5·36	—	—
265	6·22	—	—
Relative initial activity			0·67

(3) From propionate.

No. of alternate oxidations and reductions	0		1		
	Reaction velocity				
Temperature °C.	obs.	calc.	obs.	calc.	v_0/v_1 ratio
230	0·49	0·49	0·80	0·83	0·62
235	0·60	0·59	0·97	0·99	0·62
240	0·73	0·74	1·17	1·19	0·63
245	0·88	0·89	1·42	1·44	0·62
250	1·08	1·07	1·73	1·74	0·62
255	1·32	1·29	2·04	2·04	0·63
260	—	—	2·40	2·44	—
Relative initial activity					mean 0·62
Intercept on logarithmic curve		0·87		0·85	
Heat of activation		19,800		19,300 cals.	

* Owing to a slight leak these results cannot be used for temperature coefficient measurement.

Digitized by Google

TABLE III (*cont.*)

(4) From valerate.

No. of alternate oxidations and reductions	0		1		2			
	Reaction velocity						Ratios	
Temperature °C.	obs.	calc.	obs.	calc.	obs.	calc.	v_0/v_1	v_0/v_2
225	—	—	0·94	1·00	1·00	1·07	—	—
230	0·72	0·74	1·23	1·26	1·32	1·32	0·57	0·54
235	0·90	0·90	1·57	1·51	1·65	1·65	0·57	0·54
240	1·12	1·09	1·93	1·86	2·02	2·03	0·59	0·54
245	1·37	1·32	2·34	2·34	2·47	2·46	0·59	0·54
250	1·63	1·65	2·84	2·85	3·01	2·97	0·57	0·54
255	1·94	1·93	3·48	3·55	3·68	3·67	0·55	0·53
260	2·32	2·29	—	—	—	—	—	—
265	2·71	2·67	—	—	—	—	—	—
Relative initial activity							0·57	0·54
Mean							0·56	
Intercept of logarithmic curve		0·88		0·95		0·94		
Heat of activation		20,070		21,600		21,400		

Discussion of Results.

The behaviour of the catalysts produced is strikingly similar, provided that the temperature be not allowed to exceed 280° C. It is therefore tempting to conclude that the original spacing of copper atoms in the compound is only a secondary factor, that the active structure is formed by the sudden liberation of copper atoms, and that the mixed crystallisation, and freezing in the amorphous state, of such atoms is the real factor governing the nature of the final surface. Once the surface is formed it does not seem to change easily, and the activity remains steady for hours.

Very interesting also is the increase of activity on alternate oxidation and reduction. The catalyst rapidly approaches its maximum activity, after which random variation in the activity may occur to some extent, particularly if some standard procedure is not strictly adhered to in treating the catalyst while oxidising and reducing. The ratio of the initial activity to the maximum activity possible on activation has been used as a measure of the surface activity of the original preparation. With this standard it is evident that there is not much difference in the result of the decomposition of any of the simpler salts of the fatty acids.

The results therefore agree with the hypothesis that the centres of activity of the surface are frozen groups of atoms with strong specific external fields, and at moderate temperatures the stability is marked for many hours.

Digitized by Google

Note on the velocity of gas-reactions. By J. A. CHRISTIANSEN, of the University of Copenhagen. (Communicated by Mr R. H. FOWLER.)

[*Received* 10 August, *read* 25 October 1926.]

In a recent paper, G. N. Lewis and D. F. Smith* have discussed the problem of the velocity of chemical reactions, especially that of the only gas-reaction, the decomposition of N_2O_5, which had then been proved with sufficient accuracy to be kinetic unimolecular, i.e. unimolecular in the purely empiric sense of the word.

The authors try to show in this paper, that certain conclusions as to the rate with which a molecule in a gas can acquire a certain minimum amount of energy, which not only have been stated in a paper by H. A. Kramers and the author of the present note, but also before that time in somewhat different forms by several different authors†, are invalid.

These conclusions are that the number of molecules reacting per sec. in the gas (N_2O_5) as found by direct experiment, is many (at least 10^4) times greater than the number of molecules, which according to our present knowledge could possibly in the same time acquire the amount of energy which on the commonly accepted view is necessary to make the molecule react either by radiation or by collision.

Now it must be admitted that it would seem very satisfactory indeed if it could be proved that the conclusions mentioned above were invalid, but a discussion of the paper in question will show that the authors must have overlooked essential points, and that when the errors are corrected we again arrive at the original very puzzling conclusions mentioned above.

The same authors however have mentioned one circumstance, which might be of help in escaping from the difficulties mentioned, but they have not treated this point in detail. After discussion of the main points in their paper, we shall therefore, in concluding this paper, treat this point, which was to a certain degree neglected in the earlier paper by Christiansen and Kramers.

As the assumptions are explicitly stated and the train of reasoning used is set forth clearly, no report need be made here of the contents of the paper in question to which the reader is referred.

* *Journ. Amer. Chem. Soc.* **47**, p. 1508 (1925).

† *Zeitschr. Physik. Chem.* **104**, p. 451 (1923). Attention should especially be called to the remarkable papers by M. Polanyi, *Zeitschr. f. Physik*, **1**, p. 337 (1920), **2**, p. 90 (1920), **3**, p. 31 (1920).

Digitized by Google

The first point which might be discussed more deeply is the statement following the equations (13)* and (14)*:

"In this case therefore, we see that the heat of activation is not the heat required to activate the average molecule but rather the heat required to raise a molecule from zero internal energy to the condition of activation."

Now it cannot be too strongly emphasized that the "heat of activation" q can, at present, only be measured by means of experiments on the temperature coefficient of the velocity constant of the reaction. From this we calculate q by means of the well-known Arrhenius equation

$$\frac{d\ln k}{dT} = \frac{q}{RT^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

and it seems to the present author that this is at present the only possible way of getting a numerical value of this important quantity.

If we furthermore make the assumption that the probability of reaction K^* per unit time for a molecule, specified not only with respect to species but also with respect to its inner state, is strictly constant (i.e. independent of temperature and the presence of other molecules) and that the distribution of the molecules in the different possible states is that corresponding to thermal equilibrium (although the reaction is going on) we get in the unimolecular case

$$\frac{d\ln k}{dT} = \frac{q'}{RT^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2),$$

where q' denotes the difference between the average energy of the molecules which react during a certain time and the average energy of the molecules taken over all possible states, the reactive ones included†.

It is thus seen that the q found by experiment as defined in equation (1) must mean exactly the same as q' in equation (2). Evidently therefore the equation (14)* can only be a more special case of our general equation (2), a case where q' turns out to be strictly independent of temperature, which it is not in general‡.

* Equations marked * are equations so numbered in Lewis and Smith's paper.

† Comp. R. Marcelin, *Annales de Physique*, 3 [9], p. 120 (1915); R. C. Tolman, *Journ. Amer. Chem. Soc.* **42**, p. 2506 (1920). Also J. A. Christiansen, *Reaktionskinetiske Studier*, Köbenhavn, 1921. (Dissertation.)

‡ In this connection attention should be drawn once more to the paper of F. E. C. Scheffer, *Proc. Akad. Wet. Amsterdam*, **19**, p. 636 (1917), where it is proved that the integrated form of equation (1),

$$\ln k = -\frac{A}{T} + B,$$

must in all simpler cases correspond so nearly to the experimental values, that only by means of very exact experiments will it be possible to measure the deviations.

Digitized by Google

Thus contrary to the statement mentioned above the *heat of activation is always the heat required to bring an average molecule at a certain temperature to the average "state of activation" at the same temperature,* if we are justified in using the above-mentioned assumptions†.

In the special case considered by Messrs Lewis and Smith, compare their equations (11)*—(14)*, it is true that E^* is the heat required to raise a molecule from zero internal energy to the condition of activation. But E^* is *also* the energy necessary to raise an average molecule to make an average active molecule at the same temperature. The two authors must have overlooked the fact that the average energy of an active molecule, according to their definition, is not E^* but $E^* + kT$, while the average energy of all the molecules is of course kT. This can be immediately shown by means of equation (11) as

$$\bar{E}_{\text{active}} = \frac{\frac{1}{kT}\int_{E^*}^{\infty} Ee^{-\frac{E}{kT}}dE}{\frac{1}{kT}\int_{E^*}^{\infty} e^{-\frac{E}{nT}}dE} = \frac{E^*e^{-\frac{E^*}{kT}} + kTe^{-\frac{E^*}{kT}}}{e^{-\frac{E^*}{kT}}}$$

and similarly $$\bar{E}_{\text{all molecules}} = \frac{\frac{1}{kT}\int_0^{\infty} Ee^{-\frac{E}{kT}}dE}{\frac{1}{kT}\int_0^{\infty} e^{-\frac{E}{kT}}dE}.$$

The next point to be discussed is that mentioned on pp. 1514 to 1515, where it is stated that a great heat capacity of the reacting molecules would largely increase the number of collisions favourable for reaction. Now this might be true in a certain sense of the word‡, but in the paper discussed here this point is connected up very closely to the foregoing, so that it is argued that the exponent of the exponential function in the paper by Kramers and the author may be in error.

† This question is not to be discussed here, but it must be emphasized that Messrs Lewis and Smith are, in the paper considered, working on an assumption which is very similar to, although not exactly the same, as the above. Compare their equation (3) and the remark, p. 1514: "The work of Arrhenius...." The difference is, that we have replaced the right-hand side of (3),

$$K = K^* \cdot w^*,$$

by a summation (or integration), as we assume that the reacting molecules are not all in exactly the same state. This extension seems to us to be very natural and even necessary.

It may also be emphasized here that equation (2) does not tell anything of the conditions for reaction. It simply states that if the velocity constant varies with temperature there must be a difference between the mean energy of the reacting molecules and the total mean energy.

‡ This is the point, mentioned before, which will be dealt with later on.

Digitized by Google

The exponential function however cannot be much in error, as will appear from later calculations. Here again it must have been overlooked, that what we measure by experiment as the activation energy is the energy which the reacting molecules must possess in *excess* over the normal ones and consequently also the energy which must on the average be transferred by collision or by radiation at the temperature considered to any molecule to make it react. It is thus seen that the heat capacity is not relevant in connection with the exponential function.

While the two points mentioned above play only a minor rôle for the conclusions of Lewis and Smith the third point to be dealt with is a fundamental one in their paper.

In the first place the author at the time of writing (Aug. 1925) would not admit that he is convinced of the physical reality of the picture (the light-quantum hypothesis in its most paradoxical form) underlying the calculations. This however is of course no argument at all, and as it must be admitted that all consequences hitherto derived therefrom in a rational way seem to correspond to the experiments, we shall not discuss the *a priori* permissibility of using (16)* for the calculations, but shall simply take it for granted and discuss its consequences.

It seems that the authors at least partly ascribe their success to the fact that they have used the radiation hypothesis not for a narrow range of frequencies but for all frequencies exceeding a certain minimum value. This however does not help very much to make the number of activation processes great, as when ν' is great ($\sim 10^{14}$) the contributions $e^{-\frac{h\nu}{kT}} d\nu$ to the integral in (21)*

$$\phi' = \int_{\nu'}^{\infty} e^{-\frac{h\nu}{kT}} d\nu,$$

for which $\nu - \nu'$ is positive and say more than 10 °/₀ of ν' will on account of the well-known properties of the exponential function always be relatively very small at ordinary temperatures. The reason why these authors get a result of the correct order of magnitude, even in the case of a molecule of only two degrees of freedom, is that they introduce an assumption (16)* which proves to be equivalent to putting Einstein's probability for spontaneous transition from a higher to a lower quantum state (A) equal to ν, which is exactly the assumption introduced by Dushman. This however necessitates the assumption of impossible great absorption coefficients†.

The same thing is perhaps better seen by direct use of (16)*, which is believed to give the area of the cross-section of a "quan-

† Comp. Christiansen, *Zeitschr. Physik. Chem.* **103**, p. 92 (1922) and Kramers and Christiansen, *ibid.* **104**, p. 451 (1923).

Digitized by Google

tum" of definite wave-length, connected with the assumption that the "quantum" is absorbed as soon as it collides with a molecule. It ought to be mentioned that this assumption is not explicitly stated, but so far as can be seen it must be implicitly contained in the calculations if the law of conservation of energy is meant to hold, not only statistically, but also for the elementary processes, as the authors appear to assume*.

With these assumptions the cylinder described by a "quantum" on unit length (1 cm.) will have a volume of $\frac{\lambda^2}{8\pi}$. Let us take as a reasonable value $\lambda = 10^{-4}$ cm., corresponding to light in the near ultra-red, and let us shine light of this sort into a gas at 1 Atm. There will be, then, in the cylinder mentioned, about

$$2 \times 10^{19} . 10^{-8} . \frac{1}{8\pi} = 0{\cdot}1 . 10^{11}$$

molecules. If now as assumed every molecule which meets with a quantum absorbs it, then on a length of 10^{-11} cm. 1/10 of the light would be absorbed.

The absorption measured for N_2O_5† is however many million times smaller than that, and from a theoretical point of view the correspondence principle should not allow the absorption to be greater than about one millionth of the above value.

It is thus seen that the radiation theory cannot be upheld, at least if we do not make some supplementary assumptions to remove the difficulties.

Now if we were to accept the newer modification of the quantum theory, which has been proposed by Bohr, Kramers and Slater, the difficulties for the radiation theory of chemical reactions would perhaps automatically disappear, as in this theory the detailed conservation of energy is not postulated. Unfortunately, however, one of the experimental consequences has just recently been disproved by Bothe and Geiger and thus we cannot at the present time very well use this as an argument. Furthermore, quite apart from this fact, the new theory has been worked out, if I might express myself thus, so to speak in opposition to the Einstein light-quantum hypothesis which just postulates the conservation of energy, and so if we are to use the Bohr-Kramers-Slater ideas, the whole thing would have to be put in a quite different way. I might here again refer to the paper by M. Polanyi‡.

We now proceed to mention the third circumstance, which has been pointed out by Messrs Lewis and Smith, and which perhaps might allow us to escape from the difficulties.

* See for example the phrase in italics between (24) and (25).
† Warburg and Leithauser, *Ann. der Physik*, **28**, p. 313 (1909).
‡ *loc. cit.* **3**, p. 34 (1920).

Digitized by Google

As is well known there are two different ways of formally calculating the velocity of a chemical reaction.

The first one is that used by Messrs Lewis and Smith and explicitly stated in their paper. It has been restated above in a somewhat more detailed form, and has in this form been used by R. C. Tolman and others. The characteristic features of this method are the assumption of (1) equilibrium between the reactive and the normal molecules and (2) of a definite probability for reaction, corresponding to each reactive state.

The second assumption has been introduced by Marcelin. This method, which seems to be more general and which will therefore be used in the following sections, consists in enumerating the number of systems (e.g. single molecules), which *pass* a certain (or a certain number of different) state(s), the critical states, in unit time.

To use this method it is obviously necessary (1) to be able to describe these states by means of certain conditions, which they must fulfil, and (2) to calculate the number of molecules *passing* these states. Now, at the present time, our knowledge of the mechanism of any elementary process is far too limited to enable us to give a precise answer to these two questions. By means of certain simplifying assumptions it seems however possible to solve the problem for the simple case of bimolecular gas reactions, with a degree of accuracy which corresponds fairly well to the accuracy hitherto arrived at by means of experiment. The method consists for this case simply in enumerating the number of collisions between any two of the molecules, which are to react subject to the condition that their relative velocity shall exceed a certain minimum value given by the condition

$$E_0 < \frac{1}{2}\frac{m_1 m_2}{m_1 + m_2} v^2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3),$$

where m_1 and m_2 mean the masses of the two molecules. This method has been used by many different authors*, of whom especially K. F. Herzfeld ought to be mentioned, as he has applied the very essential modification of the theory, which we owe to Langevin and Rey to this case. This modification, which appeals strongly to ordinary common sense, consists, as R. C. Tolman expresses it, in taking account only of the "head-on" component of the relative velocity of the colliding molecules.

Calculated with these assumptions the velocity of reaction becomes simply

$$v = S \cdot e^{-\frac{E_0}{kT}} C_1 \cdot C_2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4),$$

where S means the ordinary gas kinetic collision number at unit

* For literature the reader is referred to *Journ. Phys. Chem.* **28**, p. 145 (1924).

Digitized by Google

concentrations. Experience shows that in all (not very numerous) cases hitherto found where we have to do with a simple bimolecular gas-reaction, this equation agrees with the experimental values, when factors greater than 1/10 and smaller than 10 are allowed for. Such small factors may be of real significance but as neither our knowledge of molecular diameters which are essential for the determination of the factor S, nor the kinetic determination of the corresponding empirical factor are very accurate, it must be stated that this method of calculation has hitherto met with a fair success*.

The picture underlying this rough estimate of the reaction velocity is the following: In order to make the two molecules react they must obtain, by the collision, a certain amount of energy *which does not depend on their internal energy.* The implicit reason for this assumption is closely connected to considerations similar to those set forth, p. 441. It is natural to assume that the molecules at the moment when they react must contain not only the critical energy E_0, as energy of parts of the molecule which are essential for the reaction, but also their normal share of thermal energy distributed on those motions in the molecule which are not essential for the reaction. Of course this is a hypothesis, but a hypothesis which seems to be supported by the results of, unfortunately rather few, experiments.

Now, however, Messrs Lewis and Smith start from another assumption. They assume that a molecule with internal energy E_i only needs to obtain the energy $E_0 - E_i$ to be activated. This extended to the bimolecular case would mean that when two molecules of internal energies E_i and E_j meet, then it is sufficient for the reaction to take place that their relative velocity V has a minimum value given by the condition

$$\frac{1}{2}\frac{m_1 m_2}{m_1 + m_2} V^2 > E_0 - (E_i + E_j) \quad \ldots\ldots\ldots\ldots\ldots (5),$$

or, to express the same thing in another way, so soon as the pair of molecules have obtained the total energy E_0 they will have at least the possibility of reacting before parting again, irrespectively of the distribution of this energy in the pair.

As the result in this case comes out very differently from that arrived at on the above-mentioned assumption, it might be of importance to use Lewis' hypothesis in a formal calculation of the reaction velocity. The current one as stated gives the simple result mentioned above.

As an expression for the number of collisions occurring per sec. per unit volume, we shall use the well-known expression from the

* In recent years C. N. Hinshelwoods' measurements of velocities of gas-reactions have contributed much to the experimental verification of equation (4).

Digitized by Google

kinetic gas-theory, given by R. H. Fowler* in the very convenient form

$$dN = 4\pi C_1 C_2 \left(\frac{m_1 m_2}{2\pi kT\,(m_1+m_2)}\right)^{\frac{3}{2}} e^{-\frac{1}{2kT}\frac{m_1 m_2}{m_1+m_2}V^2} V^3 dV \,.\, 2\pi p\, dp \qquad \text{.....(6),}$$

where p denotes the distance between the path of the centre of one molecule and the centre of the other. If now we replace in this formula C_1 and C_2 by C_i and C_j, denoting the concentrations of, respectively, molecules (1) and (2) of definite descriptions, we evidently get the number dN_A of collisions of a definite class A. If now the probability that such a collision shall be successful in giving rise to reaction is P_A, the total reaction velocity evidently becomes

$$V = \int P_A \,.\, dN_A$$

$$= 8\pi^2 \left[\frac{m_1 m_2}{(m_1+m_2)\,2\pi kT}\right]^{\frac{3}{2}} \int \ldots \int P_A \,.\, C_i \,.\, C_j\, e^{-\frac{1}{2kT}\frac{m_1 m_2}{m_1+m_2}V^2} V^3 dV \,.\, p\, dp \qquad \text{.....(7),}$$

where the integration is to be taken over all possible values of C_i, C_j, V and p. We have

$$C_i = A_1 \,.\, e^{-\frac{E_i}{kT}} d\sigma_i$$

and

$$C_j = A_2 \,.\, e^{-\frac{E_j}{kT}} d\sigma_j,$$

where $d\sigma_i$ and $d\sigma_j$ denote elements of the phase-space corresponding respectively to molecules (1) and (2), E_i and E_j the energies corresponding to the state defined by the situation of $d\sigma_i$, $d\sigma_j$ in the phase-space and A_1, A_2 are constants for given concentrations and temperature defined by means of†

$$C_1 = A_1 \int_{\text{all phases}} e^{-\frac{E_i}{kT}} d\sigma_i \qquad \text{.....(8·1),}$$

$$C_2 = A_2 \int_{\text{all phases}} e^{-\frac{E_j}{kT}} d\sigma_j \qquad \text{.....(8·2).}$$

As P_A must be a function of the coordinates specifying the collision but can be assumed to be independent of temperature, when the influence of "Einstrahlung" from the temperature radiation is not

* *Phil. Mag.* **47**, p. 257 (1924).

† A closer examination of these expressions shows that A_1 and A_2 are very closely connected with the activities of the molecules (1) and (2) respectively, and consequently

$$1/\int e^{-\frac{E}{kT}} d\sigma$$

with the activity coefficients.

Digitized by Google

taken into consideration, it is easily shown that logarithmic differentiation will lead to Arrhenius' formula for the dependence of rate of reaction on temperature (comp. p. 439), but it is obviously impossible to say anything at all of the absolute velocity without further knowledge of P_A, apart from the statement that the velocity cannot exceed the total number of collisions between (1) and (2) molecules.

If however we introduce a definite condition of the collisions being active we shall be able to use this formula for actual calculations and thus have the possibility of comparing the consequences of our assumptions with experience.

In accordance with Messrs Lewis and Smith we shall assume that $P_A = 0$ for all values of $p > \frac{1}{2}(d_1 + d_2)$, where d_1 and d_2 denote the diameters of the two molecules, and in all cases in which

$$E_i + E_j + \frac{1}{2} V^2 \frac{m_1 m_2}{m_1 + m_2} < E_0,$$

while on the other hand $P_A = 1$ when both $p < \frac{1}{2}(d_1 + d_2)$, and also

$$E_i + E_j + \frac{1}{2} V^2 \frac{m_1 m_2}{m_1 + m_2} > E_0.$$

According to Herzfeld the number of collisions between (1) and (2) with internal energies E_i and E_j which fulfil the above condition is*

$$C_i \,.\, C_j \,.\, S \,.\, e^{-\frac{E_0 - (E_i + E_j)}{kT}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (9).$$

As stated by Lewis and Smith, C_i is given, according to the classical theory, by the expression

$$C_i = C_1 \cdot \frac{E_i^{p_1 - 1}}{(p_1 - 1)!\ (kT)^{p_1}} e^{-\frac{E_i}{kT}} dE_i \quad \ldots\ldots\ldots (10{\cdot}1)$$

and similarly $$C_j = C_2 \cdot \frac{E_j^{p_2 - 1}}{(p_2 - 1)!\ (kT)^{p_2}} e^{-\frac{E_j}{kT}} dE_j \quad \ldots\ldots\ldots (10{\cdot}2),$$

where p_1 and p_2 denote integers, which can be evaluated from a knowledge of the structure of the two molecules and verified by means of measurements of the specific heat.

The reaction velocity calculated in this way consequently becomes

$$C_1 \,.\, C_2 \,.\, S e^{-\frac{E_0}{kT}} \int_0^{E_0} \frac{E_i^{p_1 - 1}}{(p_1 - 1)!\ (kT)^{p_1}} \,.\, dE_i \int_0^{E_0 - E_i} \frac{E_j^{p_2 - 1}}{(p_2 - 1)!\ (kT)^{p_2}} dE_j.$$

When we perform the integration we get

$$V = C_1 C_2 \,.\, S e^{-\frac{E_0}{kT}} \left(\frac{E_0}{kT}\right)^{p_1 + p_2} \frac{1}{(p_1 + p_2 + 1)!} \quad \ldots\ldots (11).$$

* See p. 443.

Digitized by Google

Now this expression cannot be immediately compared with an expression of the form

$$V = C_1 . C_2 . S e^{-\frac{Q}{RT}} \quad \text{.....(4),}$$

as in this S has been assumed to be independent of T. To compare, we may, for instance, differentiate (11) (for constant concentrations) logarithmically with respect to $1/T$. We thus get

$$\frac{d \ln v}{d 1/T} = -\frac{E_0}{k} + (p_1 + p_2)\, T \quad \text{.....(12),}$$

so that by comparison with Arrhenius' equation

$$\frac{d \ln v}{d 1/T} = -\frac{Q}{k},$$

$$Q = E_0 - (p_1 + p_2) kT \quad \text{.....(13)*,}$$

and

$$V = C_1 . C_2 . S e^{p_1 + p_2} e^{-\frac{E_0}{kT}} \quad \text{.....(14).}$$

To compare the expressions (11) and (14) we consequently have to calculate the ratio

$$\left(\frac{E_0}{kT}\right)^{p_1+p_2} \frac{1}{(p_1 + p_2 + 1)!} \Big/ e^{p_1+p_2}.$$

For the special case of $2N_2O_5$ molecules colliding, for which we choose with Lewis and Smith $p_1 = p_2 = 7$ and $E_0/kT = 46$, this gives a factor about 10^5.

This is to be considered as the upper limit for a factor possibly missing in equation (4) as any increase of $p_1 + p_2$ would cause a decrease of the factor, and as we have assumed $P_A = 1$ so soon as the sum of the inner energies and the relative kinetic energy exceeds E_0.

Comparison with ordinary bimolecular reactions makes it however very unlikely that this assumption is correct, as in that case we should expect to find for different reactions with different p's widely different values of the constant "S," while in fact they only vary in a rather narrow range.

On the other hand it cannot be strictly excluded that just such a special constitution of N_2O_5 would cause the peculiar exception to the general rule and we have thus to consider if the experiments might be explained on the basis of the above calculations.

Let N denote the number of "activating collisions" in a gas and w_r and w_d respectively the probabilities for reaction and de-

* This equation shows that E^* in the case considered by Lewis and Smith, *loc. cit.* p. 1519, is not 24,700 cals. ($= Q$) but $Q + 6RT = 28{,}300$ cals., so that their P^* should not be 300 but this number multiplied by $e^{-6} = 1/400$.

Digitized by Google

activation of the activated complex. Then the reaction velocity becomes

$$V = N \frac{w_r}{w_r + w_d}.$$

Now if we assume activation of (1) to take place at a collision with (2), we must necessarily on account of the reversibility of the elementary processes also assume that deactivation can take place at such a collision and we consequently get $w_d = S' C_2$, where S cannot exceed the ordinary gas kinetic collision number, while w_r denotes the probability for spontaneous reaction, which must be independent of the C's.

For N we have the ordinary expression

$$N = S'' C_1 . C_2 e^{-\frac{Q}{kT}},$$

and consequently

$$V = S'' . C_1 . e^{-\frac{Q}{kT}} \frac{w_r}{S'} \frac{1}{1 + w_r/S' C_2}.$$

Now it has been shown that at pressures as low as 1 mm. the reaction is still kinetic unimolecular and consequently

$$w_r/S' < 10^{-4} \quad \text{or} \quad w_r < 10^7.$$

Furthermore it was shown above that S'''/S' cannot exceed 10^5. Now if our picture is correct, experiment shows that $S'''/S' w_r = 10^{14}$, and consequently $w_r > 10^9$. It is thus seen that even if we press our picture as far as possible, it is impossible to avoid contradiction between the two estimates of w_r. Thus also this escape seems to be closed, and we are forced back on the hypothesis of the chain-reactions, or some other hypothesis working to the same effect. We might in this connection also mention a very peculiar hypothesis set forth by M. Polanyi*, in which a very strong coupling between two reciprocal elementary processes going on at different places in space is assumed. At the time when it was set forth, this hypothesis seemed to the present author to be exceedingly improbable, as it seemed to necessitate the conclusion that emission of radiation from a body should depend on the presence of other bodies to take up the radiation. But it must be confessed that the present development of the quantum theory is such that surprises may be expected everywhere. We may therefore for the present leave this question open and state as a result of the above considerations, that the original conclusions arrived at by M. Polanyi, Kramers and the author and others mentioned above still seem valid.

* *Zeit. f. Phys.* **3**, p. 31 (1920).

Digitized by Google

It might finally be mentioned that although it must be admitted that the experimental evidence is rather against than for a chain-reaction mechanism, it seems to the present author that the experiments, on account of serious discrepancies between the results of the different authors, are still not conclusive against some mechanism analogous to that proposed by Kramers and the author.

The author wishes also to take this opportunity to express his thanks to Mr R. H. Fowler, Trinity College, Cambridge, for the interest he has taken in the preparation of this paper.

Digitized by Google

Improvements in the apparatus for measuring the velocity of very rapid chemical reactions. II. By H. HARTRIDGE, M.D., Sc.D., Fellow of King's College, Cambridge, and F. J. W. ROUGHTON, M.A., Ph.D., Fellow of Trinity College, Cambridge.

[*Received* 12 August, *read* 25 October 1926.]

CONTENTS.

I. The measurement of reactions half completed in less than $\frac{1}{1000}$ part of a second.
II. The measurement of reactions half completed in more than one second.
III. The measurement of reactions taking place in very dilute solution.
IV. The use of smaller quantities of the reacting fluid.
V. The study of multiple reactions following one another in rapid succession.
VI. The detection of the existence of short-lived transient compounds.
VII. The detection of concentration gradients in heterogeneous reacting fluids.

During the last year the homogeneous reactions of haemoglobin, and the heterogeneous reactions of the red blood corpuscle, have continued to furnish us with abundant opportunity for the application of our method for measuring the velocity of very rapid chemical reactions*. The recent types of apparatus, which we have devised, have not departed in essentials from the general principle previously employed, but a number of modifications and extensions have been introduced to meet with greater efficiency the special conditions which were imposed upon us. Since these developments would seem to be of much wider application than the special purposes for which they were in the first place designed, we think that it may be useful to workers in other fields to give a brief summary of their general features here. In the present paper we shall accordingly describe the methods we have adopted for

1. Very fast reactions.
2. Very slow reactions.
3. Very dilute solutions.
4. Small quantities of fluid.
5. Reactions following one another in rapid succession.
6. Reactions involving short-lived transient compounds.
7. The detection of concentration gradients in heterogeneous reacting solutions.

Special apparatus has been devised to deal with each of these cases which will be briefly described in the following sections.

* Hartridge and Roughton, *Roy. Soc. Proc.* A, 104, 1923, p. 376; *Proc. Camb. Philos. Soc.* XXII, 1924, p. 426.

Digitized by Google

Section I.

The measurement of very fast reactions.

We have previously pointed out that the competency of our method to deal with fast reactions depends at present upon its ability to reduce the time of mixing, T, to as small a value as possible. This may be achieved (1) by diminishing the volume of the mixing chamber without at the same time diminishing the perfection of mixture of the fluid leaving it, (2) by increasing the rate of flow.

With regard to (1) a mere reduction in the dimensions of the mixing chamber would not prove satisfactory because the perfection of mixing would almost certainly be sacrificed thereby; it is necessary at the same time to make some alteration which will correspondingly improve the mixing. This might be (*a*) an increase in the jet velocity which would be followed by increased turbulence of the fluid (this increased velocity being achieved either by increasing the pressure in the fluids or by reducing the resistance of other parts of the apparatus); or (*b*) an increase in the number of jets feeding the reacting fluids to the mixing chamber.

In our most recent apparatus both methods were adopted. The resistance of all parts of the apparatus was kept low by making the flow of liquid to the jets as direct as possible and by reducing the length of the observation tube from 10 cms. to 3 cms. The number of jets used in the original apparatus was two, i.e. one for each of the reacting fluids. This number was increased first to four and then to eight, each time with a marked improvement in the efficiency of the apparatus. In the most recent apparatus the number was sixteen, the odd numbers being supplied with one fluid and the even numbers with the other. These jets were milled on the lathe from a solid piece of brass of conical shape. The jets thus make an angle of 60° with the axis of the observation tube. This arrangement has two advantages which amply repay the trouble expended on the construction:—in the first place it enables the mixing chamber to be reduced to much smaller dimensions and in the second it allows observation to be made on the fluids at an earlier instant after mixing. The mean distance between the commencement of the mixing chamber and the commencement of the observation tube was 2 mm., and this with a linear rate of flow of 1200 cms. per second represents an interval of time of $\frac{1}{6000}$ part of a second. The first observation can therefore be made within this very short interval of time after the commencement of mixing; and the time of half completion, R, of the fastest chemical reaction in solution which can be dealt with by means of this apparatus, is therefore just less than $\frac{1}{2000}$th of a second, i.e. a value about half that obtainable with the fastest apparatus described in our previous

Digitized by Google

paper. It seems to us improbable that we can reduce the value of R appreciably below this figure, unless either some radical change in design is made or unless the apparatus is adapted so that driving pressures much higher than one atmosphere can be applied to the reacting fluids.

Section II.

The measurement of slow reactions.

In stationary, as opposed to streaming, methods of determining the velocity of chemical reactions, it is necessary that the time taken in estimating the concentrations of the several reactants at different stages during the progress of the reaction should be short in comparison with the time taken by the reaction to proceed to half completion, i.e. R. It has been found by numerous previous

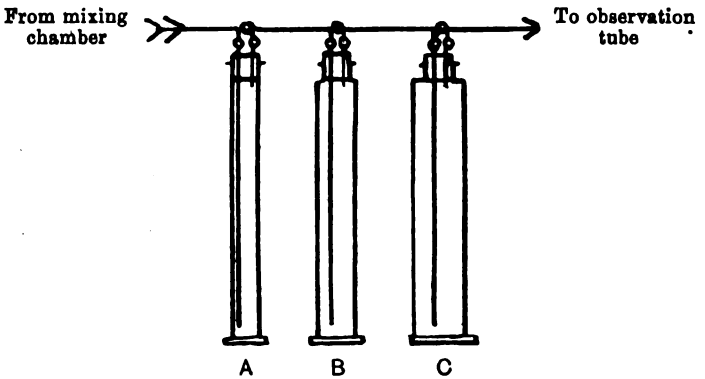


Fig. 1.

workers that no difficulty arises when R is greater than two to three minutes. But, on the other hand, as the value of R decreases from two minutes to one second, the rapidity required of the method of estimating increases correspondingly and eventually reaches an unattainable degree. We have therefore found it desirable to extend our streaming method so as to cope with reactions half-completed in times greater than one second and less than three minutes. This end has been readily achieved by simply increasing the length and hence the volume of the observation tube, and by reducing the rate of flow so far as is possible whilst still retaining turbulent motion of the fluid.

A very convenient arrangement for giving four different points on a slow velocity curve is shewn in Fig. 1.

The three cylindrical vessels and the observation tube contain ordinary air at the commencement of the experiment and the clips

Digitized by Google

on the rubber connections are so applied that fluid from the mixing chamber would first pass into *A* until it is filled, then enter and fill *B*, then enter and fill *C*, and then enter and fill an observation tube the volume of which is small in comparison with *A*, *B* and *C*, and finally flow out to waste. As soon as this occurs the first observations are taken. Then without interrupting the flow of liquid the by-pass to *C* is opened and clips to the entering and exit tubes to *C* are applied, thus cutting *C* out of the circulation. When this has been done the second series of observations are taken. Then in a similar manner *B* is next cut out and a third set of readings taken. Finally *A* is cut out and the fourth set of readings taken. From a knowledge of the rates of flow, which are taken by ascertaining the volume of liquid passing out of the apparatus in a given time for each of the four arrangements described above, and from a knowledge of the volumes of *A*, *B* and *C* and their connections the time intervals for the four points can be readily ascertained.

Section III.

The measurement of reactions taking place in very dilute solution.

It has long been recognised that it is necessary to determine the effect of wide variations in the concentration of the reagents, in order to test in a thorough manner the validity of the mechanism proposed for any given chemical reaction. Many of the physical methods of determining the composition of moving fluids, e.g. colorimetry, polarimetry and spectrometry, require a thicker volume of solution for examination when the concentration of the substance to be measured is reduced. In the case of the method of spectroscopic analysis, with which we have so largely worked, a solution which gives optimum absorption bands in an observation tube of diameter = 1 cm., required an observation tube of 50 cm. diameter, if it is diluted fifty-fold. Now it would be practically impossible to feed a tube of this diameter with fluid that had been mixed by means of a single mixing chamber. On the other hand, a number of mixing apparatus placed close together and feeding the whole cross-section of such a large tube would be very elaborate to construct.

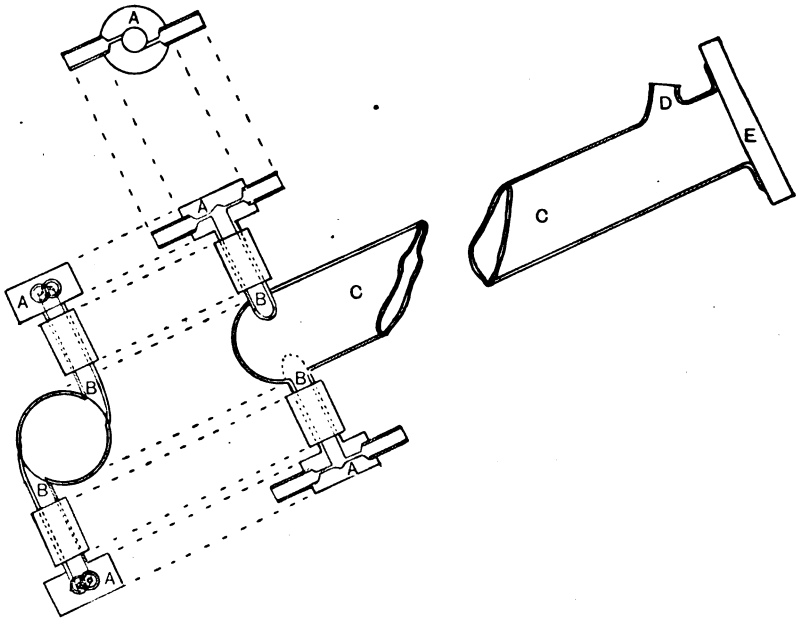
We found the difficulty could be met by placing the long axis of the observation tube in line with the optical axis of the spectroscope, as shewn in Fig. 2.

The beam of light from a Fullolite lamp placed almost in contact with the lower hemispherical end of the tube passed through the glass into the solution up which it passes. The light tends to stay inside the tube by suffering total internal reflection when it reaches glass and tries to pass out into the air. Some light gets

Digitized by Google

lost by scattering however. This loss is avoided to a large extent by covering the tube externally by a reflecting coat of good tin foil which is held in position by insulation tape wrapped round it in a spiral manner. The light having reached the top of the tube passes freely through the plate glass window which is cemented to the flange at the top and enters the revision spectroscope for analysis of the absorption bands.

The fluid passing up the observation tube has been mixed at two mixing chambers which communicate with the lower end of



A, *A* = Mixing chambers.
B, *B* = Entrances. Tubes open tangentially into observation tube.
C = Observation tube. *D* = Exit tube.
E = Plate glass cemented to flange.

Fig. 2.

the observation tube by means of two side tubes. These side tubes enter the observation tube tangentially so that a spiral motion is given to the fluid as it passes up the bore of the observation tube. This motion is given to the fluid as previously explained in order to avoid the possibility of stream line motion taking place.

The fluid having reached the top of the observation tube leaves at a side tube which is placed near the upper side of the end flange as shewn in the diagram.

Digitized by Google

Now it is very important in using this apparatus to see that the rate of flow is sufficiently rapid to preclude in the first place the possibility of stream-line motion taking place in the fluid and in the second place the possibility of the fluid entering and the fluid leaving being so different in chemical composition (owing to the reaction which is taking place in the fluid as it passes up the tube) that error is produced in the spectroscopic readings. A rate of flow higher than the critical velocity for turbulent motion will avoid the first error. The second error will also be avoided if in addition the reaction is so slow that the two extreme values of the concentration of the component under investigation differ from one another by an amount less than twice the experimental method of observation. This will be the case, as we have shewn in a previous paper*, if throughout the whole course of the reaction

$$\frac{L_0}{v} \times \frac{dc}{dt} < \frac{2A}{100},$$

where L_0 is the length of the observation tube in this case, c is concentration of the component measured, t is time in seconds, v the velocity of the fluid up the observation tube in mms. per second, and A the percentage accuracy of the method of measurement. In order to apply this test, the relation between c and t should be determined experimentally for the reaction under investigation, and by reference to the curve so obtained it can readily be determined whether or not this condition is satisfied.

The reactions in dilute solution which we have studied in these long observation tubes have all been so slow that it has been possible either to insert volumes between the mixing chamber and the observation tube much greater than the latter (as described in Section II), or to interrupt suddenly the flow of fluid and to take readings from time to time upon the stationary fluid in the observation tube. In order that sufficient light should reach the analysing instrument it is essential that the fluid should be free from all traces of turbidity, and that the ratio of the diameter of the observation tube to its length should not be less than 1 to 60. In actual practice we have worked successfully with three observation tubes of length 20 cms., 40 cms. and 160 cms. respectively. With the latter, observations on haemoglobin solutions of concentration as low as $\frac{M}{6 \times 10^6}$ have been made possible.

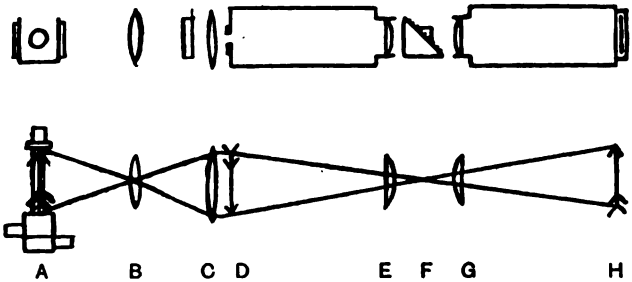
* *Proc. Roy. Soc.* A, 104, 1923, p. 388.

Digitized by Google

Section IV.

The measurement of reaction velocity on small quantities of fluid.

The apparatus which we have used previously necessitates the use of considerable quantities of fluid. As a general rule from 5 to 10 litres of each reacting fluid is necessary for an experiment. For various reasons such a quantity may be impracticable. For example one reagent may be very costly or very difficult to obtain or to prepare, or there may be a limit to the amount that is obtainable. It is the latter difficulty that is met with in our own work when it is required to measure the reaction velocities for the haemoglobin solutions prepared from the blood of a single individual, of which in the case of man or one of the small animals only about one-tenth the quantity usually required for an experiment is available.



A = Observation tube.
B = Lens for projecting image of observation tube on slit.
C = Field lens.
D = Image of observation tube at slit.
E = Lens for focusing *D* on prism.
F = Prism.
G = Lens for focusing image at prism on photographic plate.
H = Photographic plate.

Fig. 3.

Now in a given case no effective economy can usually be arranged either by reducing the rate of flow of the solutions through the apparatus or by making few observations or by making each observation shorter. Some complete change in the method is necessary, such for example as the simultaneous measurement of the reaction by four observers by means of four observation apparatus pointed at different points of the observation tube. Such an arrangement would obviously reduce the quantities of the reagents by four, and though quite practical would clearly be very inconvenient. A far better method of applying the same principle would be to have some automatic method of recording simultaneously the composition of the moving fluid at various points along the observation tube. In our particular investigation it has

Digitized by Google

been possible to photograph by means of a special spectro-camera the positions of the absorption bands at four (or more) chosen points of the observation tube, the photograph thus obtained being subsequently analysed at leisure. The principles underlying this method and some details of its employment are to be published elsewhere.

The following details of the camera may be of interest (Fig. 3). The slit was 4 mm. wide and 150 mm. long. The collimator (an aeroplane lens by Watson) had a focal length of 500 mm. and an aperture of 86 mm. The spectrum-forming system was a prism grating designed by Hartridge*. This produced a bright spectrum, the orange-yellow and green regions of which occupied the middle and lower parts of a $6\frac{1}{2}'' \times 4\frac{3}{4}''$ photographic plate. The upper part of the plate was used for recording the position of a horizontal reference line which occupied a certain fixed relationship with the spectrum. This line was produced by causing a small part of the direct beam through the collimator and camera lens to fall on the plate during its exposure to the spectrum by attaching by means of balsam the hypotenuse of a small right-angled prism to a portion of the hypotenuse of the large glass prism to which the grating was also attached. The reference line produced in this way was used to indicate the position of the spectrum on the scale attached to the densitometer which will be referred to again later.

The camera lens had a focal length of 500 mm. The dispersion of the spectrum in the yellow and green regions was approximately 22·2 A.U. per mm.

The photography of the observation tube.

By means of a large aperture projection lens of approximately 300 mm. focal length a horizontal image of the observation tube of the reaction velocity apparatus could be projected on to the slit of the spectro-camera. This image could be the same size or either reduced or increased in scale by suitably adjusting the distance between the projective lens, the observation tube and the camera. On to an end of the slit could thus be projected the part of the observation tube near the mixing chamber where the reaction has only just commenced, whereas on to the other end of the slit of the camera could be projected a part of the observation tube where the reaction was nearing completion. All the intervening portions of the slit would correspondingly receive light that had passed through the observation tube at all the other stages through which the reaction proceeds. The plate on development therefore bears a permanent record of the spectra which correspond to all these stages. In order that light may pass through the cylindrical observation tube without serious loss through refraction at its

* Hartridge, *Journ. of Physiology*, XLIX, 1915, p. 406.

Digitized by Google

surfaces the exterior of the tube is surrounded by a metal tank with parallel glass sides which is filled with water.

In order to record on the plate the time relationships of the spectra of which it bears a record a series of metal teeth a known distance apart are placed against the observation tube during the exposure of the plate so that a number of narrow lines a known distance apart are projected on to it. From a knowledge of the bore of the observation tube and the quantity of fluid passing through the tube per minute the time intervals corresponding to these marks on the plate can be ascertained. The methods of measuring the densities of the plates and their interpretation are forming the subject of separate papers to be published elsewhere.

The principle of recording the stages reached by a reaction simultaneously at different parts of the observation tube would appear to be applicable by other means beside those of photography.

Section V.

The study of multiple reactions following one another in rapid succession.

The principle underlying this method may be explained in the following way:

By means of a reaction velocity apparatus solutions of two substances *A* and *B* are caused to mix and react with one another with the production of a new substance *C*. Then if *C* is capable of reacting with a fourth substance *D* this reaction may be rapidly brought about by causing the fluid containing *C* which has just left the first reaction velocity apparatus to enter a second apparatus of similar design by which it is mixed and caused to react with *D*.

If further *C* and *D* react to produce a new substance *E*, which can react with yet another substance *F*, then the velocity of the union of *E* with *F* might be similarly studied by the use of a third reaction velocity apparatus. By thus connecting these apparatus in cascade it is possible to study multiple reactions of considerable complexity (Fig. 4).

Arrangements of this kind have proved of value for several other purposes, of which two may be mentioned here:

(1) The stage which a reaction has reached at any given instant of time can often be determined by bringing it suddenly to a stop either by the addition of some new reagent (e.g. by addition of excess of acid in the case of saponification of esters by alkali) or by a sudden change of temperature (as in the case of enzyme reactions). The concentrations reached by the different reacting substances at the instant at which the "closure" is applied can then be determined at comparative leisure. The same

Digitized by Google

principle has been made use of in some of the rapid reactions which we have studied. The principal reacting solution after leaving the first mixing chamber passes through varying intermediate volume, before it meets the "fixing" solution in the second mixing chamber, and the analysis required is then performed at comparative leisure upon the fluid issuing from the second

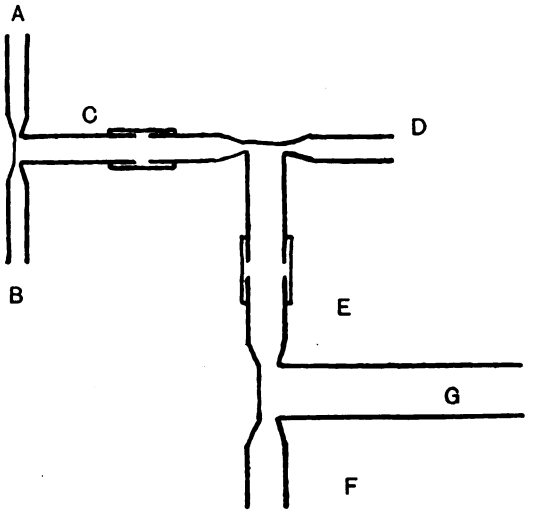


Fig. 4.

mixing chamber. The length of time during which the principal reaction was allowed to proceed can then be readily calculated from the rate of flow through the first mixing chamber and the volume traversed by the fluid between the two mixing chambers.

(2) The existence and behaviour of transient intermediate compounds can be observed in the manner explained in the next section.

Section VI.

The detection of the existence of short-lived transient compounds.

The reaction velocity apparatus provides a ready means of detecting transient compounds provided that the method of detection is applicable to the rapidly moving fluid as it travels down the observation tube of the reaction velocity apparatus. The methods of detection fall into two classes, (*a*) those in which the substance found may be detected by physical means (e.g. colour of absorption spectrum) or by physico-chemical means (e.g. *H* ion concentration) and (*b*) those in which the substance formed may be detected by chemical means, e.g. by the reaction which proceeds

Digitized by Google

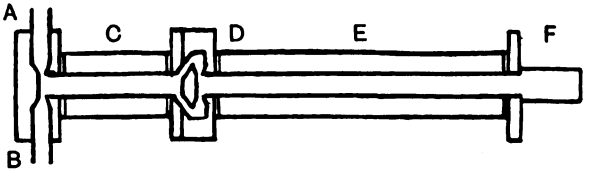
between the transient substance and some other chemical reagent which is mixed with it at an appropriate instant.

This latter method will be seen to be a special case of reactions classifiable under the previous section since it involves the employment of reaction velocity apparatus in cascade to bring about a reaction between the transient substance *C* and another reagent *D*.

Section VII.

The detection of concentration gradients in heterogeneous reactions.

In heterogeneous processes, such as the reaction between solid marble and hydrochloric acid, concentration gradients frequently arise owing to the fact that the chemical reaction proceeds rapidly at the surface of the solid, and the change in concentration produced thereby is not at once evened up throughout the fluid by



A, *B* = Entrance tubes delivering into first mixing chamber.
C = First observation tube, wherein heterogeneous reaction between *A* and *B* commences.
D = Second mixing chamber fed by bifurcation of *C*.
E = Second observation tube. *F* = Exit tube.

Fig. 5.

means of diffusion and mixing. Such gradients are usually tested for by seeing whether the observed process is quickened up by increasing the violence with which the heterogeneous fluid is stirred. Another method which we have also found useful consists in the application of a second complete and practically instantaneous mixing of the heterogeneous fluid after it has flowed far enough from the first mixing chamber (where the heterogeneous reaction is started) for a notable amount of change to have occurred. If the process shews a sudden temporary acceleration after passing through the second mixing chamber (wherein the second mixing is applied and the concentration gradients temporarily abolished), then it may be concluded that concentration gradients are tending to limit the rate observed. The apparatus required is shewn in Fig. 5.

Our thanks are due to the Medical Research Council for defraying a large portion of the expenses of this research.

Digitized by Google

Note on the Petzval optical condition. By Mr G. C. STEWARD, Gonville and Caius College.

[*Received* 13 August, *read* 25 October 1926.]

§ 1. As a preliminary to an investigation of certain diffraction patterns I was led to consider, in some detail, the geometrical aberrations of a symmetrical optical system*; and it appeared convenient then to classify the aberrations in orders according as they depend upon various powers of certain small quantities and to exhibit them as coefficients in the expansion of an 'Aberration Function.' If aberrations of the first order only are considered, it becomes evident that one of them stands, in some sense, apart from the rest; I refer to the so-called 'Petzval' condition for flatness of field. It is of interest to notice that this condition was known to Coddington and to Airy before the time of Petzval—known at least as far as concerns systems of thin lenses†. In the usual notation the condition is $\Sigma K/\mu\mu' = 0$; it is therefore independent of the positions of the object and pupil planes and in this respect it stands alone among the first order aberrations. But an increasing number of similar aberrations of higher orders will be found‡ and it is of interest to examine these and to investigate their geometrical meaning. In the following note is given a proof of the Petzval condition, differing from that usually given and falling more into line with the general theory, and indicating also a general method of examining aberrations of this peculiar type.

§ 2. *The Focal Eikonal for a Single Spherical Surface.*

Let A be the pole and O the centre of a spherical surface of radius r and power K separating media of indices μ and μ', and let F and F' be the principal foci of the surface; then

$$FA = \mu/K \text{ and } AF' = \mu'/K.$$

Let PQP' be a ray whose direction cosines are L, M, N and L', M', N' respectively, and let P and P' be the feet of perpendiculars from F and F'; let Q be upon the refracting surface.

Let U be the 'modified' focal eikonal, i.e., the eikonal multiplied by the modified power $J (\equiv K/\mu\mu')$ of the system; then, if a constant be added to ensure the vanishing of U for an axial ray, we have

$$U = J\{\mu(PQ - FA) + \mu'(QP' - AF')\} \quad \ldots\ldots\ldots (1)$$

* 'The Aberrations of a Symmetrical Optical System,' *Trans. Camb. Phil. Soc.*, vol. 23, No. 9, 1926; see also 'Aberration Diffraction Effects,' *Phil. Trans.* A, vol. 225, p. 131.

† Rayleigh, *Collected Works*, vol. 5, p. 456.

‡ T. Smith, *Trans. Opt. Soc.*, vol. 23, 1921–22.

Digitized by Google

Now $PQ = L.FO - r\cos\theta$ and $QP' = r\cos\theta' - L'.F'O$,
where θ and θ' are the angles of incidence and refraction; so that

$$PQ = \frac{\mu' r L}{\mu' - \mu} - r\cos\theta \quad \text{and} \quad QP' = r\cos\theta' + \frac{\mu r L'}{\mu' - \mu}.$$

Substituting in (1)

$$U = \frac{\mu' - \mu}{\mu\mu'}(\mu'\cos\theta' - \mu\cos\theta) + L + L' - \frac{(\mu' - \mu)^2}{\mu\mu'} - 2$$

$$= \frac{(\mu' - \mu)^2}{\mu\mu'}\left\{1 - \frac{2\mu\mu'}{(\mu' - \mu)^2}(LL' + MM' + NN' - 1)\right\}$$

$$+ L + L' - \frac{(\mu' - \mu)^2}{\mu\mu'} - 2 \ldots\ldots\ldots\ldots (2)$$

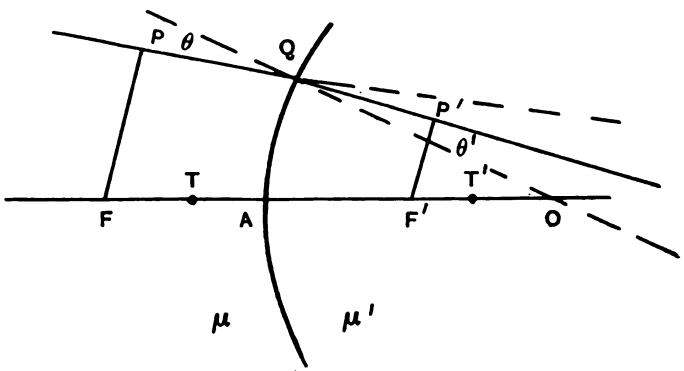


Fig. 1.

Now owing to the symmetry of the system this may be expressed in terms of the variables a, b and c given by

$$a = M^2 + N^2, \quad b = MM' + NN' \quad \text{and} \quad c = M'^2 + N'^2,$$

from which $L = \sqrt{1 - a}$ and $L' = \sqrt{1 - c}$;

then from (2), $U = \frac{1}{v}(1 + v\chi)^{\frac{1}{2}} + L + L' - \frac{1}{v} - 2, \ldots\ldots\ldots\ldots (3)$

where $v = \frac{\mu\mu'}{(\mu' - \mu)^2}$ and $\chi = -2(LL' + b - 1)$.

This expression (3) may be expanded in terms of a, b and c and the coefficients of the various terms will involve simply the function v—an invariant of the system. For our first purpose we need proceed only as far as the second order terms and to this degree of approximation we have

$$U = -b - \tfrac{1}{8}\{v(a - 2b + c)^2 + 2ca\}. \ldots\ldots\ldots\ldots (4)$$

Digitized by Google

§ 3. Let the ray intersect two conjugate planes of paraxial magnification m in points whose co-ordinates in the planes, referred to parallel rectangular axes, are Y, Z and Y', Z' respectively; let $u - \Phi$ be the modified eikonal with base points m. This may be expressed in terms of the variables α, β and γ, where

$$\alpha d^2 = (M - sM')^2 + (N - sN')^2 = a - 2sb + s^2 c,$$

$$\beta d^2 = (M - sM')(M - mM') + (N - sN')(N - mN') = a - (s + m)b + smc,$$

$$\gamma d^2 = (M - mM')^2 + (N - mN')^2 = a - 2mb + m^2 c; \ldots\ldots\ldots (1)^*$$

s is the paraxial magnification for the pupil planes and $d = s - m$. Then for perfect definition $u = f(\gamma)$ and $\Phi \equiv 0$; the form of the function $f(\gamma)$ is arbitrary and it may be determined so that the pupil planes are free from axial aberrations†. Φ may be named the 'Aberration Function' and the aberration, or wandering of the ray from the ideal focus, is given by

$$Y' - mY = \left(\frac{\partial}{\partial M'} + m\frac{\partial}{\partial M}\right)\Phi, \quad Z' - mZ = \left(\frac{\partial}{\partial N'} + m\frac{\partial}{\partial N}\right)\Phi. \qquad \ldots\ldots\ldots\ldots (2)$$

We may write $\Phi = \sum_{n=2}^{\infty} \Phi_n$ since the aberrations are essentially of order higher than the first in the variables α, β and γ; and to consider first order curvature and astigmatism we have

$$\Phi = \ldots + \tfrac{1}{8}(2\sigma_3 \alpha\gamma + 4\sigma_4 \beta^2) + \ldots,$$

from which it follows, by application of (2), that the increments in the modified curvatures of the two image surfaces produced by the system are given by

$$\Delta\left(\frac{1}{P_1}\right) = \frac{\sigma_3 + 2\sigma_4}{(s-m)^2} \text{ and } \Delta\left(\frac{1}{P_2}\right) = \frac{\sigma_3}{(s-m)^2},$$

i.e., by $\Delta\left(\frac{1}{P_1}\right) = \frac{3\sigma_4}{(s-m)^2} + \varpi$ and $\Delta\left(\frac{1}{P_2}\right) = \frac{\sigma_4}{(s-m)^2} + \varpi$,

where $\varpi (s-m)^2 = \sigma_3 - \sigma_4$. Thus $\sigma_4 = 0$ denotes the coincidence of the surfaces and then ϖ gives the measure of their common curvature.

§ 4. In fig. 1, let T and T' be the conjugate points considered; then $FT = 1/m$ and $F'T' = -m$, where the distances are modified and reduced. Thus

$$u - \Phi = U + \frac{1}{m}(1 - L) + m(1 - L'), \ldots\ldots\ldots (1)$$

* *Phil. Trans.* A, vol. 225, p. 131. † T. Smith, *loc. cit.*

Digitized by Google

and this is an identity. From § 3 (1) we have

$$4\frac{\partial^2}{\partial\gamma\partial\alpha} - \frac{\partial^2}{\partial\beta^2} = d^2\left(4\frac{\partial^2}{\partial c\,\partial a} - \frac{\partial^2}{\partial b^2}\right); \quad \ldots\ldots\ldots\ldots(2)$$

so that from (1) $\sigma_3 - \sigma_4 = d^2$,

and then $\Delta\left(\frac{1}{P}\right) = J$,

where we use reduced distances. This applies to refraction by a single spherical surface; the curvature increment produced by any number of spherical surfaces upon the same axis is therefore ΣJ and we have the general formula—valid in the absence of astigmatism—

$$\frac{1}{\mu'\rho'} - \frac{1}{\mu\rho} = \Sigma\frac{K}{\mu\mu'},$$

and the Petzval condition is

$$\Sigma\frac{K}{\mu\mu'} = 0.$$

From the above it is evident that we are dealing with transformation of the invariant type springing from § 3 (1); and further such invariants may be sought for the higher order aberrations and their precise meaning obtained from the focal eikonal for a single spherical surface given in § 2.

Digitized by Google

Maxwell's Law, and the Absorption and Emission of Radiation.

By Mr E. A. MILNE, Trinity College.

[*Received* 13 August, *read* 25 October 1926.]

1. In the famous paper* in which he developed his theory of radiation, Einstein investigated two problems. He investigated first the problem of what distribution of radiant energy in the steady state would be set up by an enclosed assembly of atoms in thermodynamic equilibrium radiating and absorbing according to the assumptions of his theory; he showed that the distribution so set up was given by Planck's law. He investigated secondly the velocities set up amongst the atoms in consequence of the random variations in direction of the emissions and absorptions; provided that emissions are assumed to be directed, so giving rise to recoil momentum, he showed that the mean square velocity-component in a given direction, $\bar{v}^2$, is equal to its equipartition value, given by $\frac{1}{2}m\bar{v}^2 = \frac{1}{2}kT$.

He did not however show that the velocity-distribution set up was Maxwellian. In view of other applications of Einstein's methods, it is proposed here to complete the investigation.

2. In a frame fixed with respect to one of the moving atoms, the radiation-field in an equilibrium enclosure is no longer isotropic. Einstein showed that, in consequence of this, an atom moving with velocity v experiences on the average a deceleration proportional† to v, say $\frac{dv}{dt} = -\lambda v$. It might at first sight be supposed that each atom would ultimately come to rest, but this is not so. Superposed on the steady deceleration there will be, in any given interval t, a net gain of velocity u, arising from chance variations in the directions of emission and absorption; u is equally likely to be positive or negative, and to the first order in v is independent of v. Following the methods he had developed in treating of Brownian motion, Einstein put as a necessary condition for a steady state

$$\overline{(ve^{-\lambda t} + u)^2} = \bar{v}^2,$$

where the bars denote averages over all the atoms.

This gives $$\bar{u}^2 = \bar{v}^2(1 - e^{-2\lambda t}),$$

or, for small t, $$\bar{v}^2 = \frac{\bar{u}^2}{2\lambda t} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1).$$

Einstein then evaluated $\bar{u}^2$ and the coefficient λ, and so determined $\bar{v}^2$.

* *Phys. Zeit.* vol. 18, p. 121, 1917.

† It is assumed that v is small compared with the velocity of light, c.

Digitized by Google

3. The theorem to be proved may now be stated as follows:

If the centre of mass of the atoms moving with velocity v moves according to the equation

$$\frac{dv}{dt} = -\lambda v \quad \text{.....(2),}$$

and if in addition in any given small interval t each atom acquires some velocity-increment u, uncorrelated with v, such that $\bar{u} = 0$, then in the steady state the velocity-distribution $f(v)\,dv$ is given by

$$f(v) = Ce^{-\frac{1}{2}v^2/\bar{v}^2} \quad \text{.....(3),}$$

where $\bar{v}^2$ is given by (1), and C is a constant. In this enunciation the word velocity means component of velocity in a given direction.

Proof. Of the atoms moving with velocity v at a given instant, let the fraction $\phi(v, w)\,dw$ acquire increments of velocity between w and $w + dw$ during the succeeding interval t. Consider the atoms which after this interval are moving with velocities between V and $V + dV$. Those of these atoms which at the beginning of the interval had a velocity v have had an increment of velocity between w and $w + dw$, where

$$w = V - v,$$

and

$$dw = dV.$$

They therefore form the fraction $\phi(v, V - v)\,dV$ of the v-atoms. But atoms moving originally with velocities between v and $v + dv$ formed the fraction $f(v)\,dv$ of the whole number of atoms. Hence if $F(V)\,dV$ is the new velocity-distribution we have

$$F(V)\,dV = dV\int_{-\infty}^{\infty} f(v)\,\phi(v,\ V - v)\,dv.$$

The condition for a steady state is $F \equiv f$, or

$$f(V) = \int_{-\infty}^{\infty} f(v)\,\phi(v,\ V - v)\,dv \quad \text{.....(4).}$$

The increment w consists of a part arising from the steady deceleration, together with an increment u. Thus

$$w = (ve^{-\lambda t} - v) + u,$$

or, for t small,

$$w = -\lambda tv + u.$$

By hypothesis, u is uncorrelated with v, and hence the function $\phi(v, -\lambda tv + u)$ is independent of v. We may therefore write it

$$\phi(v, -\lambda tv + u) = \psi(u) \quad \text{.....(5).}$$

Digitized by Google

The function $\psi(u)$ satisfies the conditions

$$\psi(u) = \psi(-u), \quad \int_{-\infty}^{\infty} \psi(u)\, du = 1,$$

and we shall write $$\int_{-\infty}^{\infty} \psi(u)\, u^2\, du = \overline{u^2}.$$

In order to make use of (5) we make a change of variable in (4). Put $v = a + bx$, and choose a and b to be constants such that identically

$$V - v \equiv -\lambda t v + x.$$

We find easily $$v = \frac{V - x}{1 - \lambda t},$$

whence $$V - v = -\lambda t \frac{V - x}{1 - \lambda t} + x,$$

and (4) becomes

$$f(V) = \int_{-\infty}^{\infty} f\left(\frac{V-x}{1-\lambda t}\right) \phi\left(\frac{V-x}{1-\lambda t}, -\lambda t \frac{V-x}{1-\lambda t} + x\right) \frac{dx}{1-\lambda t}$$

$$= \frac{1}{1-\lambda t} \int_{-\infty}^{\infty} f\left(\frac{V-x}{1-\lambda t}\right) \psi(x)\, dx \quad \ldots\ldots\ldots\ldots\ldots\ldots (6),$$

on using (5). Now expand the function f in (6) by Taylor's theorem in powers of x and integrate term by term. We find

$$f(V)(1-\lambda t) = f\left(\frac{V}{1-\lambda t}\right) + \frac{\frac{1}{2}\overline{u^2}}{(1-\lambda t)^2} f''\left(\frac{V}{1-\lambda t}\right) + \ldots$$

or $$-f(V) = \frac{1}{\lambda t}\left[f\left(\frac{V}{1-\lambda t}\right) - f(V) + \frac{\frac{1}{2}\overline{u^2}}{(1-\lambda t)^2} f''\left(\frac{V}{1-\lambda t}\right) + \ldots\right].$$

It is easily verified, either *a posteriori* or in any actual case by direct calculation, that $\overline{u^4}, \overline{u^6} \ldots$ are of a higher order in t than $\overline{u^2}$. Letting now $t \to 0$ we find

$$\tfrac{1}{2} f''(V) + [f(V) + V f'(V)] \lim_{t \to 0} \frac{\lambda t}{\overline{u^2}} = 0 \quad \ldots\ldots\ldots (7).$$

Put $$\lim_{t \to 0} \frac{\lambda t}{\overline{u^2}} = \mu.$$

The complete primitive of the differential equation (7) is

$$f(V) = e^{-\mu V^2}\left[A \int_0^V e^{\mu t^2}\, dt + B\right].$$

Since $f(V) = f(-V)$, we must have $A = 0$. Hence

$$f(V) = B e^{-\mu V^2},$$

and since $$\int_{-\infty}^{\infty} f(V)\, dV = 1,$$

we must have $$\beta = (\mu/\pi)^{\frac{1}{2}}.$$

Digitized by Google

Then $$\overline{V^2} = (\mu/\pi)^{\frac{1}{2}} \int_{-\infty}^{\infty} e^{-\mu V^2} V^2 \, dt,$$
$$= 1/2\mu.$$

Thus finally $$f(V) = \left(\frac{1}{2\pi \overline{V^2}}\right)^{\frac{1}{2}} e^{-\frac{1}{2}V^2/\overline{V^2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (8).$$

4. A differential equation equivalent to (4) has been obtained by Einstein and also by Fokker*, and is sometimes called the Einstein-Fokker equation. It has been considered in some detail by Planck†, who gave three different derivations of it. In particular, Planck emphasized the distinction between the condition for a steady state and the condition for "detailed balancing." None of these authors appear however to have given the transformations which lead from (4) to (8) in the form given above.

5. A few remarks concerning mathematical developments arising out of equation (6) may be of interest. In the first place the function ψ, though independent of v, clearly depends on t; $\overline{u^2}$, of course, depends on t. At the same time equation (4) must be identically satisfied for all t. When t is no longer restricted to be small, we must replace $1 - \lambda t$ by the exact expression $e^{-\lambda t}$, and the equation corresponding to (6) then becomes

$$f(V) = e^{\lambda t} \int_{-\infty}^{\infty} f[(V - x)\, e^{\lambda t}]\, \psi\,(x, t)\, dx \ldots\ldots\ldots (9),$$

where we have now put in evidence the dependence of ψ on t.

The left-hand side of (9) is independent of t. It can now be shown that this condition serves to determine not only f but also $\psi\,(x, t)$, i.e. to determine not only the distribution-law in the steady state but also the law of diffusion of random velocity amongst a sub-group of atoms moving initially with the same velocity.

Differentiate (9) with respect to t, and then return from the variable of integration x to the variable v by means of the relation

$$(V - x)\, e^{\lambda t} = v, \quad e^{\lambda t}\, dx = -\, dv.$$

We find

$$\lambda \int_{-\infty}^{\infty} [f(v) + v f'(v)]\, \psi \, dv + \int_{-\infty}^{\infty} f(v)\, \psi_t \, dv = 0, \ldots (10),$$

where ψ stands for $\psi\,(V - ve^{-\lambda t}, t)$ and ψ_t stands for $\partial\psi\,(x, t)/\partial t$ with $V - ve^{-\lambda t}$ replacing x after the differentiation has been performed. But from (7),

$$f(v) + v f'(v) = -\, \tfrac{1}{2} f''(v)/\mu.$$

* *Ann. d. Phys.* vol. 43, p. 810, 1914.
† *Berlin Sitz.* p. 324, 1917.

Digitized by Google

Introduce this in (10), and integrate the first integral twice by parts. We find

$$\int_{-\infty}^{\infty} \left[\frac{\frac{1}{2}\lambda e^{-2\lambda t}}{\mu} \psi_{xx} - \psi_t \right] f(v)\, dv = 0 \quad \ldots\ldots\ldots(11),$$

where ψ_{xx} denotes $\partial^2\psi(x, t)/\partial x^2$ with $V - ve^{-\lambda t}$ for x as before.

Relation (11) may be reduced to the form

$$\int_{-\infty}^{\infty} g(Ve^{\lambda t} - v, t) f(v)\, dv = 0 \ldots\ldots\ldots(11{\cdot}1),$$

which must be satisfied for all V and t. For any fixed t, this is equivalent to

$$\int_{-\infty}^{\infty} g(Ve^{\lambda t} - v) f(v)\, dv = 0 \ldots\ldots\ldots(11{\cdot}2).$$

Under certain conditions which need not be here specified, it can be shown* that this requires

$$g = 0.$$

We thus have, from (11),

$$\frac{\frac{1}{2}\lambda e^{-2\lambda t}}{\mu} \psi_{xx} = \psi_t \quad \ldots\ldots\ldots(12).$$

Putting $T = 1 - e^{-2\lambda t}$,

(12) becomes

$$\frac{1}{4\mu} \frac{\partial^2\psi}{\partial x^2} = \frac{\partial\psi}{\partial T} \quad \ldots\ldots\ldots(13),$$

the standard equation of diffusion.

The solution of (13) corresponding to a "point-source" is

$$\psi = \left(\frac{\mu}{\pi T}\right)^{\frac{1}{2}} e^{-\mu x^2/T} \ldots\ldots\ldots(14),$$

where the constant has been chosen to satisfy $\int_{-\infty}^{\infty} \psi(x)\, dx = 1$. In terms of t, this is

$$\psi(x, t) = \left(\frac{\mu}{\pi}\right)^{\frac{1}{2}} \frac{e^{\lambda t}}{(e^{2\lambda t} - 1)^{\frac{1}{2}}} e^{-\frac{\mu e^{2\lambda t} x^2}{e^{2\lambda t} - 1}} \ldots\ldots\ldots(15).$$

It is easy to see that this is in fact the solution required, and it is easily verified that (15) combined with $f(V) = (\mu/\pi)^{\frac{1}{2}} e^{-\mu V^2}$ satisfies (9) exactly for all V and t. It is also seen that

$$\overline{x^2(t)} = \int_{-\infty}^{\infty} \psi(x, t)\, x^2\, dx = \frac{e^{2\lambda t} - 1}{2\mu e^{2\lambda t}},$$

so that $\lambda t/\overline{x^2(t)} \to \mu$ as $t \to 0$, as it should.

* I am indebted to Professor Eddington for the proof of this.

Digitized by Google

As $t \to \infty$, we see from (15) that

$$\psi(x, t) \to (\mu/\pi)^{\frac{1}{2}} e^{-\mu x^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(16),$$

i.e. $\psi(x, t) \to f(x)$. The sub-group of atoms moving initially with a given velocity ultimately acquires a Maxwellian velocity-distribution which is independent of their original common velocity. Otherwise stated, each sub-group ultimately mimics the velocity-distribution of the whole assembly of which it formed a part, its centre of gravity of course coming to rest.

6. Einstein's applications of the theorem which is the subject of this paper have been to molecules in motion in an isotropic field of radiation and to particles in Brownian motion. In both these cases the system is an enclosed one in thermodynamic equilibrium. But the application of the theorem is not limited to such cases. My main object in constructing a proof of the theorem was to be able to use it in cases where there is a steady state which is not a state of thermodynamic equilibrium. Cases of this type occur frequently in the problems presented by astrophysics.

Consider a thin layer of gas exposed on one side to incident radiation. The atoms will absorb and re-emit some of this radiation, and a steady state will ultimately be set up in which as much radiation is emitted as is absorbed. This state will be characterised by a definite velocity-distribution of the atoms, and by a definite distribution of the atoms amongst the stationary states of which they are capable. But the system is not one of thermodynamic equilibrium: in the first place it is not enclosed, in the second place the radiation at any point is neither isotropic nor black. The radiation emergent on the side away from the incident radiation will in fact show absorption lines corresponding to the spectrum of the gas. We cannot therefore speak of the "temperature" of the layer of gas in the steady state. Further, we cannot apply the method of probability to determine the state of the gas, at least in the absence of theorems to that effect not yet proved. For example, we are not entitled to assume that the state of the gas is the most probable state consistent with absorbing and re-emitting so much radiation of such and such frequencies in such and such directions.

Statistical mechanics has not yet dealt to any extent with steady states in non-thermodynamic equilibrium. Of course, were the mechanisms of the reaction between the incident radiation and the atoms, and between atoms and atoms, fully known, the problem would be soluble in principle. The theorem established in the present paper enables one to short-circuit some of the analysis involved when enough is known for us to be able to assert that the conditions of the theorem are satisfied.

Digitized by Google

In the high-level atmosphere of a star, collisions between atoms are so rare as to be negligible, and the atoms are supported against downward gravity by the upward pressure of radiation. The field of radiation is non-isotropic, being confined to one hemisphere. The intensity of radiation will adjust itself so that the resulting pressure just balances gravity for an atom at rest, and chance variations in direction of emission and absorption set up a certain velocity-distribution. On an atom moving with velocity v, the difference between radiation pressure and gravity gives rise to a deceleration proportional to v, on the average, just as in Einstein's case the non-isotropy of the field to the moving atom produces a similar deceleration. We thus have the conditions postulated in the theorem. We infer that the velocity-distribution is Maxwellian, and a direct calculation of the quantities λ and $\overline{u^2}$ gives us the apparent temperature corresponding to the motions.

The details of the calculation for this case have been given by the author elsewhere*. It may be mentioned here that the result is that the radial velocity-distribution corresponds to a temperature equal to $\frac{5}{6}$ of the effective temperature of the incident radiation, whilst the cross velocity-distribution corresponds to a temperature equal to the latter temperature. The velocity-distribution in three dimensions is thus not spherically symmetrical, as in a gas obeying Maxwell's law. This further illustrates the difference between a steady state and a state of thermodynamic equilibrium.

7. Though slightly foreign to the main purpose of this note, one further point may be mentioned. It is sometimes asked in what way does Einstein's assumption of directed emissions play an essential part in his demonstration that, for atoms exposed to isotropic radiation of temperature T, we have $\frac{1}{2}m\overline{v^2} = \frac{1}{2}kT$? The answer is that the assumption affects $\overline{u^2}$ but not λ. Were the emissions spherically symmetrical waves, λ would be unaltered, but each emission would then make no contribution to $\overline{u^2}$. Since there are as many emissions as absorptions in a given interval, there would then be only half as many sudden increments of velocity, and a simple calculation shows that $\overline{u^2}$ would be halved. We should then find $\frac{1}{2}m\overline{v^2} = \frac{1}{4}kT$, showing that the velocity-distribution of the atoms under isotropic T-radiation would tend to correspond to $\frac{1}{2}T$, whilst that which they would set up under the influence of their mutual collisions would tend to correspond to T. This involves a contradiction with thermodynamics. Hence emissions must contribute recoil momentum and so must be directed.

* *Monthly Notices, R.A.S.* vol. 86, p. 578, 1926.

Digitized by Google

On compressional waves in two superposed layers. By Mr HAROLD JEFFREYS, St John's College.

[*Received* 14 August, *read* 25 October 1926.]

1. The problem discussed here arose, in a more complicated form, in some seismological work. Certain earthquakes have been found to give rise to two distinct compressional waves observable at stations up to about 600 km. away. One of these is generally regarded as a direct wave through an upper layer containing the focus, and the other as an indirect one that has travelled mainly in a deeper layer where the velocity of propagation is greater. The amplitudes of the motions produced in these two phases by a symmetrical shock were calculated on the basis of the ordinary law of refraction, the energy emitted within any pencil of rays being supposed to remain permanently within that pencil; theoretically the direct wave, at a distance of 5°, should give a movement 40 times as great as that given by the indirect one. The observed ratio is variable, but nearer to 4, and always less than 12. This shows that the ordinary law of refraction breaks down at these short distances, and makes it desirable to investigate whether a diffracted wave could give the observed motion in such circumstances. The elastic problem would, however, be very difficult, and in this paper I have only attempted the simpler one where rigidity is neglected, so that it is reduced to an application of the theory of sound. A further simplification is to neglect the curvature of the earth. With plane boundaries any ray entering the lower layer would evidently never strike the interface again, and therefore the amplitude of the indirect wave, on the simple refraction theory, would always be zero. It will be seen in what follows that this is not what actually happens.

The solution is possibly of some mathematical interest, since it is, so far as I am aware, the first instance of the application of Heaviside's methods to a case where the operational solution involves two independent differential operators.

It may be pointed out at once that irrotational disturbance will never give tangential movement at a free surface in the absence of rigidity and gravity. If Φ be the velocity potential, the condition at a free surface is that $\partial\Phi/\partial t$ is zero, and therefore Φ during the motion retains its original constant value at all points of the surface. Hence its tangential gradient is zero on the surface, and there is no tangential movement there. The horizontal movement in earthquakes is due to rigidity. Vertical movement does occur in the present case, and we may reasonably infer that its nature will be similar in the actual case.

Digitized by Google

2. Let us consider two homogeneous layers of uniform thicknesses H and K. The position coordinates are x, y, and z, the two free surfaces being $z = H$ and $z = -K$, and the interface $z = 0$ respectively. The time is t. The densities are ρ and ρ', and the velocities of compressional waves c and c', where c' is the greater. The complete velocity potentials are Φ and Φ'. Of the former, a part Φ_0 is what the impressed disturbance would produce if the medium were infinite in extent in all directions. Let us denote the operator $\partial/\partial t$ by σ, and $\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ by r^2. The condition $\Phi = 0$ at the surface $z = H$ would be satisfied automatically if Φ were equal to $\Phi_0 - \Phi_r$, where Φ_r is obtained from Φ_0 by replacing z by $2H - z$. Evidently Φ_r represents the waves reflected at the upper free surface. In addition Φ will contain a portion Φ_1 due to the other surfaces. The condition that there must be no variation of pressure over the surfaces $z = H$ and $z = -K$ then shows that Φ_1 and Φ' vanish over these surfaces.

Now Φ_1 satisfies the differential equation

$$\left(\frac{\partial^2}{\partial z^2} + r^2\right)\Phi_1 = \frac{\sigma^2}{c^2}\Phi_1 \quad \text{.....(1),}$$

and with the condition that Φ_1 vanishes at $z = H$ we see that

$$\Phi_1 = \frac{C \sinh \mu(H - z)}{\sinh \mu H} \quad \text{.....(2),}$$

where

$$\mu^2 = \frac{\sigma^2}{c^2} - r^2 \quad \text{.....(3),}$$

and C is independent of z. Similarly

$$\Phi' = \frac{D \sinh \mu'(K + z)}{\sinh \mu' K} \quad \text{.....(4),}$$

where μ' is the analogue of μ for the lower medium. At the interface $z = 0$ the pressure and the normal velocity must be continuous. If when $z = 0$

$$\Phi_0 - \Phi_r = A\,; \quad \frac{\partial}{\partial z}(\Phi_0 - \Phi_r) = B \quad \text{.....(5),}$$

these conditions give

$$\rho(C + A) = \rho' D \quad \text{.....(6),}$$

$$-\mu C \coth \mu H + B = \mu' D \coth \mu' K \quad \text{.....(7),}$$

whence

$$C = \frac{B - \mu'(\rho/\rho') A \coth \mu' K}{\mu \coth \mu H + \mu'(\rho/\rho') \coth \mu' K} \quad \text{.....(8).}$$

We require the vertical velocity at the upper surface due to Φ_1;

Digitized by Google

that due to Φ and Φ_r can easily be found directly. This velocity w is then given by

$$w = -\frac{B - \mu'(\rho/\rho')\, A \coth \mu' K}{\mu \coth \mu H + \mu'(\rho/\rho') \coth \mu' K}\, \mu \operatorname{cosech} \mu H \quad \ldots(9).$$

This is the operational solution when Φ_0 is known.

3. Let us suppose Φ_0 to consist of a spherical explosion wave, spreading out from a point on the axis of z, where $z = h (0 < h < H)$. If the pressure rises suddenly to a value proportional to $1/R$ when the wave passes a point at distance R from the focus, and thereafter remains constant, the equations of motion are satisfied; the corresponding velocity potential can be denoted symbolically by

$$\Phi_0 = -\frac{1}{\sigma R} e^{-\sigma R/c} \quad \ldots\ldots\ldots\ldots\ldots\ldots(10).$$

Bromwich's interpretation is that $f(\sigma)$ is defined to mean

$$\frac{1}{2\pi\iota} \int f(\gamma)\, e^{\gamma t}\, \frac{d\gamma}{\gamma},$$

taken along a line of infinite length parallel to the imaginary axis and on the positive side of it, subject of course to $f(\gamma)$ having such a form that this integral has a meaning. Alternatively we may replace γ by $\iota\gamma$ and integrate along a line parallel to the real axis and below it; this convention sometimes makes it easier to translate results already obtained by Fourier's theorems. Then

$$\Phi_0 = \frac{1}{2\pi} \int \frac{1}{R\gamma^2} e^{\iota\gamma(t - R/c)}\, d\gamma \quad \ldots\ldots\ldots\ldots\ldots\ldots(11).$$

The operational form of the vertical velocity at $z = H$ is

$$w_0 = \frac{\partial}{\partial z}(\Phi_0 - \Phi_r) = 2\,\frac{H - h}{cR^2}\left(1 + \frac{c}{\sigma R}\right) e^{-\sigma R/c} \ldots\ldots(12).$$

The exponential factor shows that there is no motion up to time R/c. The first term then jumps suddenly to $2(H - h)/cR^2$ and thereafter remains constant. The second increases steadily with the time after the wave arrives, but this term is less important the greater the distance. The above expression gives the motion due to the direct wave and its first reflexion at the upper boundary.

The expression for Φ_0 must now be transformed into cylindrical harmonics. If ϖ be the distance from the vertical line through the focus, ν^2 becomes the Bessel operator $\frac{1}{\varpi}\frac{\partial}{\partial \varpi}\left(\varpi \frac{\partial}{\partial \varpi}\right)$; and the operator μ^2 acting on $J_0(\lambda\gamma\varpi/c)\, e^{\iota\gamma t}$ is equivalent to the constant

$$(\gamma^2/c^2)(\lambda^2 - 1).$$

Digitized by Google

Also μ'^2 is equivalent to $(\gamma^2/c^2)(\lambda^2 - c^2/c'^2)$. Let us define μ and μ' to be positive when λ is real and greater than 1, and suppose a cut in the λ-plane so that λ moves round the branch-points 1 and c/c' on the positive imaginary side. Then when λ is real and positive, and either μ or μ' is imaginary, the latter is positive imaginary. Now when $\varpi = 0$

$$\frac{1}{R} e^{-\iota\gamma R/c} = \frac{1}{\zeta} e^{-\iota\gamma\zeta/c} \quad (\text{where} \quad \zeta = |z - h|)$$

$$= \int_{\iota\gamma/c}^{\infty} e^{-\mu\zeta} d\mu$$

$$= \frac{\gamma}{c} \int_0^{\infty} e^{-\mu\zeta} \frac{\lambda d\lambda}{\sqrt{(\lambda^2 - 1)}} \quad \ldots\ldots\ldots\ldots\ldots\ldots(13),$$

and hence when ϖ is not zero

$$\frac{1}{R} e^{-\iota\gamma R/c} = \frac{\gamma}{c} \int_0^{\infty} e^{-\mu\zeta} J_0(\lambda\gamma\varpi/c) \frac{\lambda d\lambda}{\sqrt{(\lambda^2 - 1)}} \quad \ldots\ldots(14).$$

This may also be obtained by considering the values of the function on the left when ζ is zero and using the Fourier-Bessel integral theorem. Sommerfeld* obtains a function analogous to that on the right as a representation of $\frac{1}{R} e^{\iota\gamma R/c}$. We need now the values of Φ_0 and $\partial\Phi_0/\partial z$ when ζ is h; finding these from (14) and substituting in (9) we have for the remaining portion of the vertical velocity

$$w_1 = -\frac{1}{2\pi} \iint_0^{\infty} \frac{1}{c\gamma} \exp(\iota\gamma t - \mu h) J_0(\lambda\gamma\varpi/c)$$

$$\times \frac{\mu - \mu'(\rho/\rho') \coth \mu' K}{\mu \coth \mu H + \mu'(\rho/\rho') \coth \mu' K} \mu \operatorname{cosech} \mu H \frac{d\gamma d\lambda}{\sqrt{\lambda^2 - 1}} \ldots(15),$$

together with a term derived from Φ_r, which can easily be written down when required.

In the paper just mentioned Sommerfeld treats a problem of the propagation of electromagnetic waves in two media with a plane interface, which has many points of resemblance to the present one. He however takes both media as of unlimited vertical extent, whereas the finite depth of the upper medium is essential to the seismological case. There are also practical advantages in many wave problems in taking the media to be of finite extent, especially when operational methods of solution are used. For instance, $\coth \mu' K$ in the above solution is an analytic function; but if we put K infinite before solution, $\coth \mu' K$ would have to be replaced by $+1$ when μ' has a positive real part, and by -1

* *Ann. d. Physik*, vol. 28, 1909.

Digitized by Google

when it has a negative real part, so that the convenience of working with an analytic function would be lost. Sommerfeld's problem, again, treats an infinite train of harmonic waves, so that he is restricted to a single value of γ; here we are concerned with a single pulse. He is able to evaluate his single integral with regard to λ by contour integration; but the variation with γ of the positions of the poles would introduce difficulties in evaluating the above double integral.

4. An alternative mode of attack is suggested by Bromwich's expansion method*. In interpreting an operator with, say, $\sinh \sigma T$ in its denominator, we can expand in descending powers of $e^{-\sigma T}$ as if this were a small quantity, and then interpret term by term; each term then gives a separate pulse. The process is equivalent to replacing the line integral with regard to γ by one around a large semicircle, on which the expansion is easily justified; and then each term of the expansion in a series of integrals is equivalent to a corresponding integral along the original straight path. In the present case, if $\lambda > 1$, μ and μ' are real and positive, and the justification is immediate; if $\lambda < 1$ Bromwich's process establishes it. In our representation we can then put

$$\frac{\mu - \mu'(\rho/\rho')\coth \mu' K}{\mu \coth \mu H + \mu'(\rho/\rho')\coth \mu' K} \operatorname{cosech} \mu H$$

$$= \frac{\mu \sinh \mu' K - \mu'(\rho/\rho')\cosh \mu' K}{\mu \cosh \mu H \sinh \mu' K + \mu'(\rho/\rho')\sinh \mu H \cosh \mu' K}$$

$$= 2e^{-\mu H}\frac{\mu - \mu'\rho/\rho'}{\mu + \mu'\rho/\rho'}\left(1 + \frac{\mu - \mu'\rho/\rho'}{\mu + \mu'\rho/\rho'}e^{-2\mu H} + \text{terms in } e^{-2\mu' K}\right)^{-1}$$

.....(16).

The first term of this expression will give the most direct waves of the set here considered. The term in $e^{-3\mu H}$ represents a wave that has undergone two additional reflexions, one at the upper surface and one at the interface, and so on for the higher powers of $e^{-\mu H}$ that arise in the expansion. Terms in $e^{-2\mu' K}$ and higher powers represent pulses reflected at the lower boundary. We are concerned mainly with pulses that have been reflected not more than once at the interface, and need therefore consider only the first term. Thus (15) is reduced to

$$w_1 = -\frac{1}{\pi}\iint_0^\infty \frac{1}{c^2}\exp\{\iota\gamma t - \mu(H+h)\}\frac{\mu - \mu'\rho/\rho'}{\mu + \mu'\rho/\rho'}J_0(\lambda\gamma\varpi/c)\,\lambda\, d\gamma\, d\lambda$$

.....(17).

Now we may express the Bessel function in terms of the two

* *Proc. Lond. Math. Soc.* vol. 15, 1916, pp. 425–430.

Digitized by Google

Hankel integrals, the first of which vanishes when the imaginary part of the argument is $+\iota\infty$, and the second when it is $-\iota\infty$. Then

$$2J_0 = H_0^{(1)} + H_0^{(2)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(18).$$

The integral (17) with respect to λ can thus be broken up into two, one involving $H_0^{(1)}$ and the other $H_0^{(2)}$. The former is equivalent to an integral to $\iota\infty$, the contour being completed by an integral around a large quadrant. The integrand is everywhere regular within this contour. The latter integral can similarly be replaced by one to $-\iota\infty$ and one round a large arc in the fourth quadrant; but with our sign conventions there are branch-points where μ and μ' vanish, so that the contour must be completed by loops from the infinite arc drawn so as to exclude these points. There is no pole within the contour (here the problem differs from that treated by Sommerfeld). The two integrals along the imaginary axis cancel, as in Sommerfeld's problem, and those around the infinite arcs tend to zero in the limit. The integral with regard to λ is then equivalent to two integrals along narrow loops around the branch-points, say Q and Q', of the form

$$\int \exp\{-\mu(H+h)\}\frac{\mu-\mu'\rho/\rho'}{\mu+\mu'\rho/\rho'}H_0^{(2)}(\lambda\gamma\varpi/c)\,\lambda\,d\lambda \quad \ldots(19).$$

If the loops are so drawn that their sides are nearly parallel to the imaginary axis, and $\lambda\gamma\varpi/c$ is great, the important parts of the integrals along them are concentrated in the neighbourhoods of the branch-points. When ϖ is great we may approximate to the integrals by substituting for the Hankel function the first term of its asymptotic expansion, thus

$$H_0^{(2)}(x) \sim \sqrt{\frac{2}{\pi x}}\exp\{-\iota(x-\tfrac{1}{4}\pi)\} \quad \ldots\ldots\ldots(20).$$

Let us write $$c/c' = \alpha;\ (1-\alpha^2)^{\frac{1}{2}} = \beta \quad \ldots\ldots\ldots\ldots\ldots\ldots(21).$$

5. Considering first Q, the integral (19) around the branch-point $\lambda = 1$, put

$$\lambda = 1 - \iota\theta \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(22),$$

where θ is real and positive. The integrand is negligible except where θ is small. Then

$$\mu = \frac{\gamma}{c}\sqrt{(2\theta)}\,e^{-\frac{1}{4}\pi\iota};\ \mu' = \beta\gamma/c \quad \ldots\ldots\ldots\ldots\ldots(23)$$

approximately, on the part of the path going from the branch-point to $-\iota\infty$; on the other side the sign of μ must be reversed.

Digitized by Google

Making these substitutions, we have

$$Q = -2\left(\frac{2c}{\pi\gamma\varpi}\right)^{\frac{1}{2}} \exp\left\{-\iota\left(\frac{\gamma\varpi}{c} - \tfrac{1}{4}\pi\right)\right\} \int_0^\infty e^{-\gamma\varpi\theta/c} \times \left[\sinh\mu(H+h) + \frac{2\mu\rho'}{\mu'\rho}\cosh\mu(H+h)\right] \iota d\theta \ldots\ldots(24),$$

where terms involving higher powers of μ have been neglected in the coefficients of the hyperbolic functions. But

$$\int_0^\infty e^{-k\theta}\sinh m\sqrt{\theta}\, d\theta = \tfrac{1}{2} m \sqrt{\frac{\pi}{k^3}} \exp\frac{m^2}{4k} \ldots\ldots(25),$$

whence the portion of Q arising from $\sinh\mu(H+h)$ is found to be

$$-\frac{2\iota(H+h)c}{\gamma\varpi^2}\exp\left\{-\frac{\iota\gamma}{c}\left(\varpi + \frac{(H+h)^2}{2\varpi}\right)\right\} \ldots\ldots(26).$$

The corresponding portion of the vertical velocity is

$$w_2 = -\frac{1}{2\pi c}\int \frac{2\iota(H+h)}{\gamma\varpi^2}\exp\left[\frac{\iota\gamma}{c}\left(ct - \varpi - \frac{(H+h)^2}{2\varpi}\right)\right] d\gamma$$

$$= \frac{2(H+h)}{c\varpi^2} e^{-\sigma R_2/c} \ldots\ldots(27),$$

to our order, where R_2 is the distance of the point considered from the image of the focus in the interface. This represents the wave reflected in the interface. Comparing this with (12) we see that it gives more movement than the direct wave.

The portion of Q arising from $\cosh\mu(H+h)$ can be found by differentiating (26) with regard to $H+h$ and introducing a factor $2\rho' c/\rho\gamma\beta$. It gives two new waves arriving at the same time as that just discussed; one of these is of the same type, but its amplitude contains the insignificant factor $(H+h)^2/\varpi^2$; the other gives a velocity starting from zero and increasing steadily with the time, like the second term in the expression for the direct wave. Neither of these appears important, except that they represent the loss of energy from the original wave due to transmission into the lower medium.

6. Considering Q' next, we put

$$\lambda = \alpha - \iota\theta;\ \mu = \iota\beta\gamma/c;\ \mu' = (\gamma/c)(2\alpha\theta)^{\frac{1}{2}} e^{-\frac{1}{4}\pi\iota} \ldots\ldots(28).$$

On the two sides of the loop μ has the same value, but the two values of $\dfrac{\mu - \mu'\rho/\rho'}{\mu + \mu'\rho/\rho'}$ differ by approximately $\dfrac{4\mu'\rho}{\mu\rho'}$. Then we find

$$Q' = \frac{4\rho}{\rho'}\frac{c^2}{\gamma^2\varpi^2}\frac{\alpha}{\beta}\exp\left[-\frac{\iota\gamma}{c}\{\alpha\varpi + \beta(H+h)\}\right] \ldots\ldots(29),$$

Digitized by Google

and the corresponding part of the vertical velocity is, in operational form,

$$w_3 = \frac{4\rho}{\rho'}\frac{\alpha}{\beta\varpi^2}\frac{1}{\sigma}\exp\left[-\frac{\sigma}{c}\{\alpha\varpi + \beta(H+h)\}\right] \quad \ldots\ldots(30).$$

This represents a wave arriving at time $\frac{\varpi}{c'} + \frac{\beta}{c}(H+h)$. If the times of arrival at different places on the upper surface are compared, the wave will appear to travel at a rate c'; its apparent velocity is therefore the velocity of sound in the lower layer. But the term depending on $H+h$ implies a delay in arriving equal to that expected on the ordinary law of refraction owing to the time spent in the upward and downward passages through the upper layer. The ordinary law of refraction therefore gives the correct time of arrival of the indirect wave. But the velocity does not jump to a finite value when the wave arrives, as for the direct wave and that reflected at the interface; it begins at zero and steadily increases with the time. As for the waves in the upper layer, the velocity contains a factor ϖ^{-2}, so that the decrease of the amplitude with increasing distance, which was the chief difficulty of the law of refraction, may be considered explained.

The type of wave here considered is not one likely to occur in practice, but others can be derived from it easily. An important case is that of a wave spreading out from a sphere initially at a uniform pressure above that of its surroundings, all the fluid being initially at rest. In such a case the primitive velocity potential is of the form

$$\Phi_0 = -\frac{1}{2\sigma^2 R}\left[e^{-\frac{\sigma}{c}(R-a)}\left(a\frac{\sigma}{c}-1\right) + e^{-\frac{\sigma}{c}(R+a)}\left(a\frac{\sigma}{c}+1\right)\right] \quad \ldots\ldots(31),$$

where a is the radius of the sphere. The results for this case are therefore to be obtained by operating on the previous ones with

$$\frac{1}{2\sigma}\left[\left(\frac{a\sigma}{c}-1\right)e^{\sigma a/c} + \left(\frac{a\sigma}{c}+1\right)e^{-\sigma a/c}\right] \quad \ldots\ldots(32).$$

Considering first w_0, the velocity due to the direct wave, we have the operational solution

$$\frac{H-h}{cR^2}\left(1+\frac{c}{\sigma R}\right)\left[\left(\frac{a}{c}-\frac{1}{\sigma}\right)e^{-\sigma(R-a)/c} + \left(\frac{a}{c}+\frac{1}{\sigma}\right)e^{-\sigma(R+a)/c}\right] \quad \ldots\ldots(33).$$

The first exponential gives zero up to time $(R-a)/c$, the second till $(R+a)/c$. When t lies between these critical values

$$w_0 = \frac{H-h}{cR^2}\left[\frac{a}{c} - \left(1-\frac{a}{R}\right)\tau - \frac{1}{2}\frac{c}{R}\tau^2\right] \quad \ldots\ldots(34),$$

Digitized by Google

where τ is the time since the arrival of the first wave. After the time when $\tau = 2a/c$ the velocity again vanishes. During the passage of the wave the velocity changes steadily from $+\frac{a(H-h)}{c^2R^2}$ to $-\frac{a(H-h)}{c^2R^2}$. Surface particles do not return quite to their original level; after the passage of the wave there is a small residual elevation $\frac{2}{3}\frac{a^3(H-h)}{c^3R^3}$. The maximum displacement is nearly at the middle of the passage and amounts to $\frac{1}{2}\frac{a^2(H-h)}{c^2R^2}$ nearly.

The corresponding results for the wave reflected at the interface are found by changing the sign of h.

The velocity due to the indirect wave is, in operational form,

$$w_3 = \frac{2\rho\alpha}{\rho'\beta\varpi^2}\left[\left(\frac{a}{c\sigma}-\frac{1}{\sigma^2}\right)\exp\left\{-\frac{\sigma}{c}\{\alpha\varpi+\beta(H+h)-a\}\right\}\right.$$
$$\left.+\left(\frac{a}{c\sigma}+\frac{1}{\sigma^2}\right)\exp\left\{-\frac{\sigma}{c}\{\alpha\varpi+\beta(H+h)+a\}\right\}\right] \quad ...(35).$$

There is no motion up to time $\{\alpha\varpi+\beta(H+h)-a\}/c$, and the velocity is zero after time $2a/c$ later. If we denote the time from the beginning of the movement by τ as before, the velocity while the movement lasts is

$$w_3 = \frac{2\rho\alpha}{\rho'\beta\varpi^2}\left(\frac{a}{c}\tau - \tfrac{1}{2}\tau^2\right) \quad .....(36).$$

The velocity starts at zero, increases to a maximum, and decreases to zero again at the rear of the wave. The displacement therefore increases throughout the passage of the wave. Its final and maximum value is $\frac{4}{3}\frac{\rho\alpha a^3}{\rho'\beta\varpi^2c^3}$. After the wave has passed a permanent upward displacement remains. The ratio of the maximum displacements in the indirect and direct waves is $\frac{8}{3}\frac{\rho\alpha}{\rho'\beta}\frac{a}{H-h}$. This number is independent of ϖ. In the seismological analogue ρ/ρ' is about $\frac{4}{5}$, and α/β is nearly unity. Thus the ratio is nearly $2a/(H-h)$. The ratio of the maximum displacements in the direct and indirect waves therefore depends only on the ratio of the linear dimensions of the focus to the depth of focus.

The direct and indirect waves therefore produce decidedly different effects in their passage. In the former the arrival of the wave gives instantly a finite velocity; in the latter it gives a finite acceleration. The maximum displacement in the former case is at a time nearly equal to a/c after the beginning of the disturbance;

Digitized by Google

in the latter the corresponding interval is $2a/c$. Finally the direct wave leaves no appreciable residual displacement afterwards; the indirect one leaves a displacement equal to the maximum during its passage.

In the seismological case rigidity would modify these results slightly; it would prevent the persistence of a displacement after the passage of the wave, for such a displacement would imply a distortion, and would therefore be annihilated later by the generation of new distortional waves. It seems unlikely that rigidity would modify the other relations to any important extent.

7. *Summary.* The assumption that the compressional waves of earthquakes follow the ordinary laws of refraction, the energy within any pencil of rays remaining permanently within that pencil, has been found to lead to too small amplitudes for the indirect waves from near earthquakes. In this paper a system consisting of two superposed compressible, but non-rigid, media is considered; the lower is supposed to transmit compressional waves with the greater velocity. It is found that an explosion within the upper medium produces a disturbance at the upper surface involving the direct wave and all the reflected waves that might be expected; but in addition a wave is found that appears to have travelled along the interface with the velocity of sound in the lower medium. This indirect wave would have zero amplitude on the simple laws of refraction with plane boundaries. The variation of its amplitude with distance from the focus is in reasonable accordance with seismological observation, and its time of arrival agrees with that inferred from the laws of refraction for boundaries with slight curvature. But if the direct wave begins with a finite velocity, the indirect one will begin with a finite acceleration; in seismological language an iP_g will be associated with an eP_n. The indirect wave will also take longer than the direct one to give its maximum displacement.

Digitized by Google

PROCEEDINGS AT THE MEETINGS HELD DURING THE SESSION 1925—1926.

ANNUAL GENERAL MEETING.

October 26, 1925.

In the Cavendish Laboratory,

PROFESSOR J. T. WILSON, PRESIDENT, IN THE CHAIR.

The following were elected Officers for the ensuing year:

President:

Professor J. T. Wilson.

Vice-Presidents:

Mr C. T. Heycock.
Dr G. T. Bennett.
Mr G. Udny Yule.

Treasurer:

Mr F. A. Potts.

Secretaries:

Mr J. Gray.
Mr F. P. White.
Mr R. H. Fowler.

Members of the Council who continue:

Mr H. McCombie.
Mr J. B. S. Haldane.
Dr A. B. Appleton.
Dr C. D. Ellis.
Mr F. F. Blackman.
Prof. C. T. R. Wilson.
Mr J. E. Littlewood.

New Members of Council:

Prof. Nuttall.
Dr J. Chadwick.
Mr H. Munro Fox.
Mr W. B. R. King.
Mr P. M. S. Blackett.

The following was elected a Fellow of the Society:

V. C. Robinson, B.A., Gonville and Caius College.

The following were elected Associates of the Society:

E. N. Allott, Emmanuel College.
G. Bateson, St John's College.
Miss A. M. Bidder, Newnham College.
E. S. Castle, Trinity Hall.

Digitized by Google

Miss K. M. Chamberlain, Girton College.
G. Eadie, Sidney Sussex College.
B. A. Edwards, Emmanuel College.
I. J. Faulkner, Emmanuel College.
Dr S. Hecht.
Miss P. M. Jenkin.
T. C. Mitchell, Corpus Christi College.
D. M. Morrison, Trinity College.
Miss M. E. Robinson, Girton College.
Miss B. Swirles, Girton College.
E. M. Taylor, St John's College.
L. G. Vedy, Downing College.
L. Wertenstein.
T. R. Wilkins.
W. R. Wright.
C. E. Wynn-Williams, Trinity College.

The following Communications were made to the Society:

1. The natural X-ray spectrum of radium B. By Professor Sir E. RUTHERFORD and W. A. WOOSTER, B.A., Peterhouse.

2. (1) The atomic number of a radioactive element at the moment of emission of the γ-rays.
(2) The β-ray type of disintegration.
By Dr C. D. ELLIS, Trinity College, and W. A. WOOSTER, B.A., Peterhouse.

3. The analysis of the β-ray spectrum due to the natural L-radiation of radium B. By D. H. BLACK. (Communicated by Professor Sir E. Rutherford.)

4. Doublet and triplet separations in optical spectra as evidence whether orbits penetrate into the core. By Mr D. R. HARTREE, St John's College.

5. On the reciprocation of one quadric into another. By Professor H. F. BAKER.

6. The enumeration of the partitions of multipartite numbers. By Major P. A. MACMAHON.

7. Assemblies of imperfect gases by the method of partition functions. By Mr R. H. FOWLER, Trinity College.

8. Note on the pedal locus. By Mr J. P. GABBATT, Peterhouse.

9. An extended form of Kronecker's theorem, with an application which shows that Burgers' theorem on adiabatic invariants is statistically true for an assembly. By L. H. THOMAS, B.A., Trinity College. (Communicated by Mr R. H. Fowler.)

10. On the series of forms of Jacobi's rotating liquid ellipsoid. By Mr W. F. SEDGWICK, Trinity College.

Digitized by Google

November 9, 1925.

In the Botany School,

PROFESSOR J. T. WILSON, PRESIDENT, IN THE CHAIR.

The following were elected Fellows of the Society:

R. W. Ditchburn, B.A., Trinity College.
A. B. Miller, M.Sc., Trinity College.

Professor B. M. Jones delivered a Lecture entitled "Wings."

November 23, 1925.

In the Comparative Anatomy Lecture Room,

PROFESSOR J. T. WILSON, PRESIDENT, IN THE CHAIR.

The following were elected Fellows of the Society:

P. J. Durrant, B.A., Selwyn College.
S. Goldstein, B.A., St John's College.

The following were elected Associates of the Society:

J. G. Hare, Sidney Sussex College.
G. I. Hoover, Emmanuel College.
W. S. Hughes, Fitzwilliam Hall.
C. H. N. Jackson, Gonville and Caius College.
B. Lewis, Fitzwilliam Hall.
E. G. Lewis, Emmanuel College.

The following Communications were made to the Society:

1. On a case of matriclinous inheritance in *Antirrhinum*. By Professor R. C. PUNNETT.

2. The change with age of a linkage between Mendelian factors in fowls. By Mr J. B. S. HALDANE, Trinity College.

3. A note on the inheritance of yellow fat in rabbits. By Mr M. S. PEASE, Trinity College.

4. Notes on the electrical counter. By K. EMELÉUS, B.A., St John's College. (Communicated by Mr E. V. Appleton.)

December 7, 1925.

In the Cavendish Laboratory,

PROFESSOR J. T. WILSON, PRESIDENT, IN THE CHAIR.

The following was elected an Associate of the Society:

Miss C. A. Scott, Girton College.

Digitized by Google

The following Communications were made to the Society:

1. The quenching of resonance radiation and the breadths of spectral lines. By R. W. DITCHBURN, B.A., Trinity College.

2. A photographic method of determining the mobility of recoil atoms. By G. H. BRIGGS. (Communicated by Dr J. Chadwick.)

3. The effect of the crystal boundaries on the deformation of aluminium crystals. By R. L. ASTON. (Communicated by Mr G. I. Taylor.)

4. The adiabatic hypothesis for magnetic fields. By P. A. M. DIRAC. (Communicated by Mr R. H. Fowler.)

5. (1) Periodic solutions of a linear differential equation of the second order with periodic coefficients.
(2) The second solution of the Mathieu equation.
By Mr E. L. INCE, Trinity College.

6. The equianharmonic and harmonic envelopes of a quadrinodal cubic surface. By G. TIMMS. (Communicated by Mr W. P. Milne.)

7. On the substitution of Wallis's postulate of similarity for Euclid's postulate of parallels. Addendum. By Dr M. J. M. HILL, Peterhouse.

8. On a family of constructs in higher space. By C. G. F. JAMES, B.A., Trinity College.

January 18, 1926.

In the Cavendish Laboratory,

PROFESSOR J. T. WILSON, PRESIDENT, IN THE CHAIR.

The following was elected a Fellow of the Society:

D. Burnett, B.A., Clare College.

The following were elected Associates of the Society:

C. W. Chamberlain.
H. M. Lewis, Corpus Christi College.
J. R. Oppenheimer, Christ's College.

The following Communications were made to the Society:

1. On chains of two-two relations, and the theory of elliptic functions. By Prof. H. F. BAKER.

2. The aberrations of a symmetrical optical system. By Mr G. C. STEWARD, Emmanuel College.

3. On a group of 1440 birational transformations of four variables that arises in considering the projective equivalence of double sixes. By Mr W. BURNSIDE, Pembroke College.

4. Overtensions in a condenser battery during a sudden discharge. By Dr P. L. KAPITZA, Trinity College.

Digitized by Google

5. Photocatalysis at mercury surfaces. By H. S. HIRST. (Communicated by Mr E. K. Rideal.)

6. The period of decay of RaB and RaC. By P. BRACELIN. (Communicated by Dr J. Chadwick.)

February 1, 1926.

In the Chemical Laboratory,

MR C. T. HEYCOCK, VICE-PRESIDENT, IN THE CHAIR.

The following was elected a Fellow of the Society:

C. T. L. Caton, B.A., Trinity College.

The following was elected an Associate of the Society:

Alex. Ree.

The following Communications were made to the Society:

1. Metallic complexes with aliphatic poly-amines. By Dr F. G. MANN, Downing College, and Professor Sir W. J. POPE.

2. Isoquinoline and the isoquinoline reds. By Mr R. T. M. HAINES, Gonville and Caius College, and Professor Sir W. J. POPE.

3. The behaviour of the centres of activity of saturated surfaces during the initial stages of unimolecular reaction. By F. H. CONSTABLE, B.A., St John's College.

4. Direct titration of thallous salts by potassium iodate. By Mr A. J. BERRY, Downing College.

5. Influence of surfaces on the combination of ethylene with the halogens. By Dr R. G. W. NORRISH, Emmanuel College.

6. The chlorination and bromination of 4-amino-diphenyl. By W. A. WATERS. (Communicated by Mr H. McCombie.)

7. Oxidation at charcoal surfaces. By Miss M. WRIGHT. (Communicated by Mr E. K. Rideal.)

February 15, 1926.

In the Botany School,

PROFESSOR J. T. WILSON, PRESIDENT, IN THE CHAIR.

The following was elected a Fellow of the Society:

H. L. H. H. Green, M.A., Sidney Sussex College.

The following were elected Associates of the Society:

H. I. Coombs, Trinity College.
Miss N. F. Paterson, Newnham College.

Professor A. V. Hill delivered a Lecture entitled "Irreversible phenomena in muscular shortening."

Digitized by Google

March 1, 1926.

In the Comparative Anatomy Lecture Room,

PROFESSOR J. T. WILSON, PRESIDENT, IN THE CHAIR.

The following were elected Fellows of the Society:

A. E. Watkins, M.A., St John's College.
S. W. Watson, B.A., Gonville and Caius College.

The following Communications were made to the Society:

1. The development of a harpacticid copepod. By Mr F. A. POTTS, Trinity Hall.

2. The growth of fish during the phase of yolk sac nutrition. By Mr J. GRAY, King's College.

3. Intracellular oxidation-reduction potential and anaerobiosis. By Dr J. NEEDHAM, Gonville and Caius College.

4. Insular gravity and oceanic isostasy. By Professor Sir J. LARMOR.

5. (1) Relativity in connexion with axial rotation.
(2) The problem of relativity in reference to several bodies.
By Mr R. HARGREAVES, St John's College.

6. The general (m, n) correspondence. By R. VAIDYANATHASWAMY. (Communicated by Mr H. W. Turnbull.)

7. The stability of a strut under thrust, when buckling is resisted by a force proportional to the displacement. By S. GOLDSTEIN, B.A., St John's College.

March 15, 1926.

In the Cavendish Laboratory,

PROFESSOR J. T. WILSON, PRESIDENT, IN THE CHAIR.

The following were elected Fellows of the Society:

C. A. Meredith, B.A., Trinity College.
W. A. D. Rudge, M.A., St John's College.

The following were elected Associates of the Society:

M. C. Henderson, Trinity College.
Hon. Mrs Huia Onslow, Newnham College.
Mrs Veitch, Girton College.

Professor A. Sommerfeld delivered a Lecture entitled "Some controversial points in the theory of spectra."

The following Communication was made to the Society:

A theory of the action of X-rays on living cells. By Dr J. A. CROWTHER, St John's College.

Digitized by Google

May 3, 1926.

In the Cavendish Laboratory,

PROFESSOR J. T. WILSON, PRESIDENT, IN THE CHAIR.

The following Communications were made to the Society:

1. Insular gravity and oceanic isostasy. By Professor Sir J. LARMOR.

2. A mechanical model of the Rutherford-Bohr atom. By Mr W. A. D. RUDGE, St John's College.

3. On the higher singularities of plane algebraic curves. By Miss C. A. SCOTT. (Communicated by Mr F. P. White.)

4. Sur la non-existence simultanée de deux fonctions de Mathieu. By Ž. MARKOVIĆ. (Communicated by Mr E. L. Ince.)

5. On the multiple tangents and multisecants of scrolls in higher space. By C. G. F. JAMES, B.A., Trinity College.

6. On the diurnal variation of ultra-short wave wireless transmission. By Mr E. V. APPLETON, St John's College.

7. The (2, 1) Correspondence. By R. VAIDYANATHASWAMY. (Communicated by Mr H. W. Turnbull.)

8. On the attenuation of wireless waves in short distance overland transmission. By J. A. RATCLIFFE, B.A., Sidney Sussex College, and M. A. F. BARNETT. (Communicated by Mr E. V. Appleton.)

July 26, 1926.

In the Cavendish Laboratory,

MR G. UDNY YULE, VICE-PRESIDENT, IN THE CHAIR.

The following were elected Fellows of the Society:

H. Graham Cannon, M.A., Christ's College.
B. W. Robinson, B.A., Trinity College.
W. E. Woodward, M.A., Trinity College.

The following were elected Associates of the Society:

A. Caress, Trinity Hall.
Lieut.-Col. J. E. Craster.
Miss E. D. Pepper.

The following Communications were made to the Society:

1. On quantum algebra. By Dr P. A. M. DIRAC, St John's College. (Communicated by Mr R. H. Fowler.)

2. The polarisabilities of atomic cores. By Miss B. SWIRLES, Girton College. (Communicated by Mr R. H. Fowler.)

3. On the quantum theory of the problem of the two bodies. (Preliminary communication.) By J. R. OPPENHEIMER. (Communicated by Mr R. H. Fowler.)

Digitized by Google

4. Integral invariants of the affine field. By Mr M. H. A. NEWMAN, St John's College.

5. The linear invariants of ten quaternary quadrics. By Dr A. YOUNG, Clare College, and Mr H. W. TURNBULL, Trinity College.

6. A contribution to the theory of Ferromagnetism. By G. S. MAHAJANI, B.A., St John's College.

7. On some resistance properties of a certain net-work containing inductances and capacities, and their analogies in a mechanical system. By Mr E. B. MOULLIN, Downing College. (Communicated by Dr G. F. C. Searle.)

8. On Mellin's inversion formula. By Mr J. C. BURKILL, Trinity College.

9. The elliptic products of Jacobi and the theory of linear congruences. By Major P. A. MACMAHON.

10. Geodetic and dynamical principles, a comparison and connexion. By Mr R. HARGREAVES, St John's College.

11. On the quantum theory of vibration-rotation bands. By J. R. OPPENHEIMER. (Communicated by Mr R. H. Fowler.)

12. An approximation to the motion of two rotating electrical doublets in a plane. By P. A. TAYLOR, B.A., Emmanuel College.

13. Some relations between the optical spectra of different atoms of the same electronic structure. II. Aluminium-like and copper-like atoms. By Mr D. R. HARTREE, St John's College.

14. Note on the extension to higher space of a theorem of Wallace. By Mr J. P. GABBATT, Peterhouse.

15. A mathematical theory of natural and artificial selection. Part III. By Mr J. B. S. HALDANE, Trinity College.

16. An optical illusion due to contrast. By Mr G. P. THOMSON, Corpus Christi College.

Digitized by Google

STATEMENT OF ACCOUNTS OF THE CAMBRIDGE PHILOSOPHICAL SOCIETY, FROM JANUARY 1, 1925, TO DECEMBER 31, 1925.

Receipts.

	£	s.	d.	£	s.	d.
Balance due to the Society, Jan. 1, 1925	34	16	2			
With Treasurer on petty cash account	2	6	3			
				37	2	5
Subscriptions paid in advance	3	3	0			
Arrears of Subscriptions	10	10	0			
Subscriptions for 1925	473	12	6			
Associates' Subscriptions for 1925 ...	77	3	6			
Tickets for Library	18	10	0			
				582	19	0
Dividends on Liverpool Corporation Stock	8	9	4			
Dividends on Conversion Loan	26	7	10			
				34	17	2
Sale of Publications				15	0	3
Interest on Deposit Account for 1925 ...				17	10	0
Royal Society: Publication Grant ...				150	0	0
				837	8	10
Balance due from the Society, Dec. 31, 1925	48	14	4			
Less in hand: On petty cash account ...	1	7	11			
				47	6	5
				£884	15	3

Expenditure.

	£	s.	d.	£	s.	d.
University Press	42	8	6			
„ „	5	13	6			
				48	2	0
„ „ (Clay)	665	17	0			
„ „ („)	158	5	4			
				824	2	4
Geological Society of America				3	7	7
Cheque Book...					10	0
Carriage of parcels and postage etc. ...				8	13	4
				£884	15	3

SUSPENSE ACCOUNT.

Receipts.

	£	s.	d.
Balance due to the account Jan. 1, 1925...	300	0	0
Composition Fee	21	0	0
University of Durham: Sale of Publications (Back Numbers)	35	12	0

Expenditure.

	£	s.	d.
Purchase of £376. 0*s.* 9*d.* Conversion Loan @ 3½ per cent.	300	0	0
Balance	56	12	0

Receipts.

	£	s.	d.
Balance due to the Fund, Jan. 1, 1925	329	9	11
Dividends on £750 L.N.E.R. 4 per cent. Debentures:			
January, 1925	11	12	6
July, 1925	11	16	3
	£352	18	8

Expenditure.

	£	s.	d.
Balance due to the Fund, Dec. 31, 1925	352	18	8
	£352	18	8

STATEMENT OF BALANCES.

Receipts.

	£	s.	d.	£	s.	d.
Balance at Messrs Barclays, Dec. 31, 1925:						
On Deposit	739	12	2			
Less due from the Society on current account overdrawn	378	15	10			
				£360	16	4

Expenditure.

	£	s.	d.	£	s.	d.
Balance due to Hopkins Prize Fund Account	352	18	8			
Balance due to Society on Suspense Account	56	12	0			
	409	10	8			
Less due from Society on Current Account	48	14	4			
				£360	16	4

CAPITAL ACCOUNT (RESERVE FUND).

	£	s.	d.
£307. 1*s.* 10*d.* Liverpool Corporation 3½ per cent. Stock ...	225	6	10
£957. 7*s.* 8*d.* Conversion Loan 3½ per cent. @ 76¼	730	1	10
(Present value)	£955	8	8

F. A. POTTS, *Treasurer*

REPORT OF THE AUDITORS.

We have examined the Treasurer's accounts and find them correct. We have also verified the investments of £307. 1*s.* 10*d.* Liverpool poration Stock, £957. 7*s.* 8*d.* Conversion Loan, £750 L.N.E.R. Co. 4 per cent. Debenture Stock.

J. T. SAUNDERS
E. C. FRANCIS } *Auditors.*

February 17th, 1926.

Digitized by Google

Digitized by Google

PROCEEDINGS
OF THE
Cambridge Philosophical Society.

Geodetic and Dynamical Principles, a comparison and connexion.

By Mr R. HARGREAVES, St John's College.

[*Received* 8 June, *read* 26 July 1926.]

§ 1. The geodetic problem presented by a quadratic form is identical with the dynamical problem in which the same quadratic appears in the kinetic energy, and potential energy is not represented. If the quadratic in space-differentials is equated to $k dt^2$, the problem is essentially the same whatever constant value is given to k; the corresponding feature is the arbitrary constant value of kinetic energy.

If we define ds^2 by

$$ds^2 = (C + 2U)\,dt^2 - (a_{11}\,dx_1^2 + 2a_{12}\,dx_1\,dx_2 + \ldots) \quad \ldots(1),$$

where U and the coefficients a depend on the variables x, then one way of dealing with the geodetic problem is to write

$$\int ds = \int \left[(C + 2U) \left(\frac{dt}{ds}\right)^2 - \left\{ a_{11} \left(\frac{dx_1}{ds}\right)^2 + \ldots \right\} \right] ds$$

and apply the dynamical method with all total differential coefficients taken with regard to s in lieu of the usual t. A second method consists in writing

$$\int ds = \int \{C + 2U - (a_{11}\dot{x}_1^2 + 2a_{12}\dot{x}_1\dot{x}_2 + \ldots)\}^{\frac{1}{2}}\, dt \quad \ldots\ldots(2),$$

where $\dot{x} = dx/dt$. In the selection of one variable as independent, possible difficulties might arise if the variable selected had a limited range, which will not occur when the choice falls on an ignorable coordinate. In the problem of relativity t is ignorable, and the natural variable to choose; it is also in (1) taken to be reversible.

The point gained in working through (2) is an immediate expression in terms of t as independent variable, but it is at the cost of working with a kinetic potential L, not itself normal, but a function of a normal kinetic potential $\lambda = T - U$. The treatment is based on the identity

$$\frac{d}{dt}\left(\Sigma \dot{x} \frac{\partial L}{\partial \dot{x}} - L\right) = \Sigma \left(\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x}\right) \dot{x} \quad \ldots\ldots\ldots\ldots(3),$$

Digitized by Google

which is true for L, any function of $x\ldots$ and $\dot{x}\ldots$. When L is a function of λ, the equations of motion have the form

$$\frac{d}{dt}\left(\frac{dL}{d\lambda}\frac{\partial\lambda}{\partial\dot{x}}\right)-\frac{dL}{d\lambda}\frac{\partial\lambda}{\partial x}=0 \quad \ldots\ldots\ldots\ldots\ldots\ldots(4),$$

and the equation of energy is

$$E=\Sigma\dot{x}\frac{\partial L}{\partial\dot{x}}-L=2T\frac{dL}{d\lambda}-L \quad \ldots\ldots\ldots\ldots\ldots\ldots(4').$$

In (2) we are concerned with $L=-k(C-2\lambda)^{\frac{1}{2}}$, where the constant k may be useful in adjusting dimensions for the dynamical analogue. With this value of L the equation of energy is

$$E=\frac{k(C+2U)}{(C-2\lambda)^{\frac{1}{2}}} \quad \ldots\ldots\ldots\ldots\ldots\ldots(5),$$

and an equation of motion is

$$\frac{d}{dt}\frac{k}{(C-2\lambda)^{\frac{1}{2}}}\frac{\partial T}{\partial\dot{x}}=\frac{k}{(C-2\lambda)^{\frac{1}{2}}}\frac{\partial\lambda}{\partial\dot{x}} \quad \ldots\ldots\ldots\ldots\ldots\ldots(6).$$

In (6) we are concerned on the right hand only with the value, not with the mode of expression of $k/(C-2\lambda)^{\frac{1}{2}}$, on the left hand also with its complete time-rate, but not with any partial differential coefficients. Hence on each side $k/(C-2\lambda)^{\frac{1}{2}}$ may be replaced by its value $E/(C+2U)$, with a result

$$\begin{aligned}\frac{d}{dt}\frac{E}{C+2U}\frac{\partial T}{\partial\dot{x}}&=\frac{E}{C+2U}\left(\frac{\partial T}{\partial x}-\frac{\partial U}{\partial x}\right)\\&=\frac{\partial}{\partial x}\frac{ET}{C+2U}+\frac{E(2T-C-2U)}{(C+2U)^2}\frac{\partial U}{\partial x}\\&=\frac{\partial}{\partial x}\frac{ET}{C+2U}-\frac{E(C-2\lambda)}{(C+2U)^2}\frac{\partial U}{\partial x}\\&=\frac{\partial}{\partial x}\frac{ET}{C+2U}-\frac{k^2}{E}\frac{\partial U}{\partial x} \quad \ldots\ldots\ldots\ldots\ldots\ldots(7).\end{aligned}$$

These are equations of a normal dynamical system in which

$T'=ET/(C+2U)$ is kinetic, and $U'=k^2U/E$ is potential energy
$\ldots\ldots\ldots\ldots\ldots\ldots(8).$

Also (5) may be written as

$$E(C-2T+2U)/(C+2U)=k^2(C+2U)/E,$$

or

$$\tfrac{1}{2}\left(E-\frac{k^2C}{E}\right)=\frac{ET}{C+2U}+\frac{k^2U}{E}=T'+U' \quad \ldots\ldots\ldots(9);$$

which gives the energy constant to be used with (8) to secure absolute agreement with the working of the original form. This completes the transformation of a problem stated in geodetic

Digitized by Google

terms to a normal problem of dynamics. It is a variant of a theorem given in an earlier paper*, of which the basis was an application of the principle of ignoration. It is hoped that the restatement is justified by some addition to the clearness and significance of the theorem.

§ 2. An easy example is the path of a ray of light, where we are concerned with a stationary value of Vt as given by a quadratic form $V^2dt^2 = \gamma^{-2}dr^2 + \gamma^{-1}r^2d\phi^2$, where $V^2\gamma = V^2 - 2m/r$. Here we take the ignorable variable ϕ as independent variable, write $\dot{r} = dr/d\phi$ and the problem is that of $\int L d\phi$, where

$$L = -(\gamma^{-2}\dot{r}^2 + \gamma^{-1}r^2)^{\frac{1}{2}}.$$

Thus
$$E = \dot{r}\frac{\partial L}{\partial \dot{r}} - L = \frac{\gamma^{-1}r^2}{(\gamma^{-2}\dot{r}^2 + \gamma^{-1}r^2)^{\frac{1}{2}}},$$

or
$$\frac{1}{E^2} = \frac{1}{r^4}\left(\frac{dr}{d\phi}\right)^2 + \frac{\gamma}{r^2} = \left(\frac{du}{d\phi}\right)^2 + \gamma u^2 \quad \ldots\ldots\ldots\ldots(10).$$

The orbit of a planet as determined by a geodetic method applied to Einstein's quadratic falls under the general method, with $k = V$, $C = V^2$, $U = -(m_1 + m_2)/r = -m/r$ say,

$$\gamma V^2 = V^2 + 2U, \quad 2T = \gamma^{-1}\dot{r}^2 + r^2\dot{\phi}^2.$$

In the new system

$$2T' = E(\gamma^{-2}\dot{r}^2 + \gamma^{-1}r^2\dot{\phi}^2)/V^2, \quad U' = -mV^2/Er \quad \ldots(11).$$

Thus with $H = Er^2\dot{\phi}/\gamma V^2$,

$$\left(\frac{du}{d\phi}\right)^2 + \gamma u^2 = \frac{1}{H^2}\left(2mu + \frac{E^2}{V^2} - V^2\right) \quad \ldots\ldots\ldots(12)$$

determines the orbit.

* *Proc. Camb. Phil. Soc.*, vol. 22, p. 138 (1924), esp. p. 156.

Digitized by Google

On the Figure of Pappus' Theorem. By Mr C. V. HANUMANTA RAO, Trinity College. (Communicated by Professor H. F. BAKER.)

[*Received* 20 October, *read* 25 October 1926.]

When a triangle ABC is in perspective with $A'B'C'$ and with $B'C'A'$, so is it also in perspective with $C'A'B'$, and the two triangles may then be said to be in cyclic perspective. The centres of perspective then form a third triangle XYZ, such that, of the three triangles, every two are in cyclic perspective with the vertices of the third for centres. The figure thus arising is the figure of Pappus' Theorem with nine points and nine lines incident by threes.

We here consider the configuration of Pappus as consisting of the vertices of three triangles, every two of which are in cyclic perspective, together with the nine lines on which they lie by threes. When two triangles are in perspective, it is known that there exists a conic in regard to which the one triangle reciprocates into the other, and the centre of perspective is the pole of the axis. There thus arise nine conics *associated* with the nine points respectively: the conic with respect to which ABC reciprocates into $A'B'C'$ is thus associated with the centre X. It is a fact that these nine conics are in three groups of three, each group consisting of three conics of which any one reciprocates in regard to any second into the third: the conics associated with A, B, C form one such group. It is also a fact that the nine conics can be arranged in nine sets of three, each set being a set of three conics which are linearly connected: the conics associated with A, A', X form one such set.

Further, the nine conics are found to be capable of being expressed as linear functions of three of them which are not themselves linearly related. It is, in general, true that three conics which are linearly independent are the polar conics of a unique set of three points in regard to a unique cubic curve. An exception arises, as pointed out by Rosanes*, in the case where a linear function of the three conics can be found to reduce to a pair of coincident lines. Such an exception arises in the present case; nevertheless *there exists a unique cubic curve with respect to which the nine conics are the polar conics of the nine centres*, and this indicates the manner of association of the conics and the centres spoken of above. Moreover, this cubic curve is *equianharmonic*, and its Hessian consists of three lines forming a triangle LMN which is self conjugate in regard to each of the nine conics. Of the

* *Math. Ann.*, vol. XVII, 1880, p. 21.

Digitized by Google

elements of this triangle only one vertex L and the opposite side MN are real; of the three pairs of vertex and opposite side, and in particular of the vertex L and the opposite side MN, it is true that the line is the polar of the point* in regard to each of the three triangles in the figure. The triangle LMN may thus be said to be self polar with respect to each of the three triangles started with; such a relation of self polarity of one triangle in regard to a second is a mutual relation; and conversely it is true to say that when two triangles are both self polar in regard to a single further triangle, the two triangles are necessarily in cyclic perspective, and so is the third triangle which is in cyclic perspective with both of them also self polar in regard to the further triangle mentioned. It can also be shown that when a point and a line are pole and polar for a triangle and also for a conic, so are they similarly related in regard to the reciprocal of the triangle with respect to the conic.

The point L is the point of concurrence of the three real inflexional tangents of the cubic; and these three lines through L form a cubic form which is simultaneously apolar with the cubic forms arising from the joins of L to A, A', X, the joins of L to B, B', Y, and the joins of L to C, C', Z; and similar statements can be made in regard to M and in regard to N.

Again, a conic as a geometric figure has two expressions, one in point coordinates, and one in line coordinates. From the latter point of view, one may speak, with Clebsch, of a conic as the polar conic of a line in regard to a class cubic. It is then true *that the nine conics are the polar conics of nine lines in regard to a certain class cubic*, each conic being associated with a point as above, and with a line as here, such that the point and the line are a vertex and the opposite side of one of the three triangles. For instance, the conic for which ABC reciprocates into $A'B'C'$ is associated with the point X and with the line YZ. This class cubic has a Hessian, viz. the Jacobian of its polar conics, consisting of the three points L, M, N themselves. All the results in this note can be verified with the help of appropriate analysis.

* A point and a line are said to be pole and polar with respect to a triangle when the projections of the point from the vertices on the opposite sides are the harmonic conjugates in regard to the vertices on the side considered of the intersection of that side with the line. Cf. Baker, *Principles of Geometry*, vol. II, p. 38.

Digitized by Google

On a group of order 25920 *and the projective transformations of a cubic surface.* By Mr W. BURNSIDE, Pembroke College.

[*Received* 2 November, *read* 22 November 1926.]

The figure in three dimensions consisting of two sets of three non-intersecting lines, such that each line of one set intersects each line of the other set, will here be called a "net." All nets are projectively equivalent to each other: and there are just 72 projective transformations that change one net into another.

Every cubic surface that contains the net N, or

$$\left.\begin{matrix} y=0 \\ z=0 \end{matrix}\right\} \quad \left.\begin{matrix} y-t=0 \\ z-x=0 \end{matrix}\right\} \quad \left.\begin{matrix} t=0 \\ x=0 \end{matrix}\right\} \quad \left.\begin{matrix} y=0 \\ t=0 \end{matrix}\right\} \quad \left.\begin{matrix} y-z=0 \\ t-x=0 \end{matrix}\right\} \quad \left.\begin{matrix} z=0 \\ x=0 \end{matrix}\right\}$$

is given by

$$K[xy(z+t)-zt(x+y)]+(Ax+By+Cz+Dt)[xy-zt]=0;$$

and therefore every irreducible cubic surface that contains N is given by

$$xy(z+t)-zt(x+y)+(Ax+By+Cz+Dt)(xy-zt)=0,$$

for some finite values of A, B, C and D.

If Schläfli's notation is used for the 27 lines on the surface, and if the six lines of N in order are denoted by a, b, c, d', e', f', then the equations of the nine lines ad, ae, af, bd, be, bf, cd, ce, cf can be written rationally in terms of A, B, C, D, [e.g. the equations of ad are $y=0$, $(A+1)x+Cz+Dt=0$] and the equations of the remaining twelve lines can be written rationally in terms of A, B, C, D and E, where

$$E^2=\tfrac{1}{4}(1+A+B+C+D+AB+CD)^2-ABCD.$$

Let N' be any net on $S=0$, or

$$xy(z+t)-zt(x+y)+(Ax+By+Cz+Dt)(xy-zt)=0,$$

and let T' be any one of the 72 projective transformations that change N' into N. Then, since the equations of the lines of N' are rational in A, B, C, D, E, the coefficients in the transformation T' are also rational in A, B, C, D, E; and, therefore, the coefficients in the cubic surface $T'S=0$ are rational in A, B, C, D, E. But the surface $T'S=0$ contains the net $N=0$. Hence the equation $T'S=0$ is

$$xy(z+t)-zt(x+y)+(A'x+B'y+C'z+D't)(xy-zt)=0,$$

where A', B', C', D' are rational in A, B, C, D, E.

Digitized by Google

Now N' may be any one of the 360 nets on $S=0$, and T' may therefore be any one of 360×72 $(=25920)$ projective transformations. There are therefore just 25920 surfaces

$$xy(z+t)-zt(x+y)+(A'x+B'y+C'z+D't)(xy-zt)=0,$$

which are projectively equivalent to $S=0$ and therefore to one another; and the set arises from any one of them by the same process by which they arise from $S=0$. Thus, if

$$E'^2=\tfrac{1}{4}(1+A'+B'+C'+D'+A'B'+C'D')^2-A'B'C'D',$$

then A, B, C, D are rational functions of A', B', C', D', E'; just as A', B', C', D' are rational functions of A, B, C, D, E. Hence, if between the equations expressing these relations we eliminate either A', B', C', D (or A, B, C, D), we find that $E'=$ rational function of A, B, C, D, E (or $E=$ rational function of A', B', C', D', E'). We are thus led to the set of relations

$$A'=f_1(A,B,C,D,E),\ B'=f_2(A,B,C,D,E),\ C'=f_3(A,B,C,D,E),$$
$$D'=f_4(A,B,C,D,E),\ E'=f_5(A,B,C,D,E)$$

where f_1, f_2, f_3, f_4, f_5 are rational functions; and where A, B, C, D, E are rational functions of A', B', C', D', E'.

Now if from the first four of the equations their values are written for A', B', C', D' in

$$E'^2=\tfrac{1}{4}(1+A'+B'+C'+D'+A'B'+C'D')^2-A'B'C'D'$$

there results

$$E'^2=g(A,B,C,D,E)$$

where g is a rational function; and this can only be consistent with

$$E'=f_5(A,B,C,D,E)$$

if $g(A,B,C,D,E)$ is the square of a rational function of A, B, C, D, E. Hence the process indicated sets up, from the 25920 projective transformations, 2×25920 birational transformations of the five symbols A, B, C, D, E;

$$A'=f_1(A,B,C,D,E),\ B'=f_2(\quad),\ C'=f_3(\quad),\ D'=f_4(\quad),$$
$$E'=\pm f_5(A,B,C,D,E),$$

being the two which arise from one projective transformation. These form a group; and in respect of the group of order two, constituted by

$$A'=A,\quad B'=B,\quad C'=C,\quad D'=D,\quad E'=E,$$
$$A'=A,\quad B'=B,\quad C'=C,\quad D'=D,\quad E'=-E,$$

it is simply isomorphic with a group of order 25920.

This group of order 25920 thus arises directly in connection with the projective transformations of a cubic surface which has just 27 distinct lines on it.

Digitized by Google

The Compton Effect in Wave Mechanics. By Dr P. A. M. DIRAC, St John's College.

[*Received* 8 November, *read* 22 November 1926.]

§ 1. *Introduction.*

The problem of the scattering of radiation by a free electron has been treated by the author* on the basis of Heisenberg's matrix mechanics, which was first modified to be in agreement with the principle of relativity. The main point of this modification is that, whereas in the non-relativity theory one deals with matrices whose elements vary with the time according to the law $e^{i\omega t}$, in the relativity theory the elements of the matrices must vary according to the law $e^{i\omega t'}$ where $t' = t - (l_1 x_1 + l_2 x_2 + l_3 x_3)/c$ if they are to determine correctly the radiation emitted in the direction specified by the direction cosines (l_1, l_2, l_3), x_1, x_2 and x_3 being the coordinates of the electron at the time t. These matrices were obtained by writing the Hamiltonian equation of the system in the form

$$H' - W' = 0 \quad \text{.....(1),}$$

where W' is a variable canonically conjugate to t' and H' commutes with t', and then using H' as an ordinary Hamiltonian function of a dynamical system that has W' for its energy and t' for its time variable.

This method is rather artificial; particularly so since on the quantum theory there is no unique way of writing the Hamiltonian equation in the form (1) [owing to ambiguity in the order of factors in H'], and it has to be proved that all reasonable ways of doing so lead to the same results. A more natural and more easily understood method of obtaining the matrices is provided by Schrödinger's wave mechanics†.

In this method we do not need to write the Hamiltonian equation in any special form, such as (1), in order to obtain Schrödinger's wave equation from it, by applying the rules which consist of letting the momentum symbols denote certain operators. When we have obtained the solutions ψ_a of the wave equation, then we can get a matrix representation of the dynamical variables by obtaining expansions of the type

$$A\psi_a = \Sigma_{a'} A_{a'a} \psi_{a'},$$

* *Roy. Soc. Proc.* A, vol. 111, p. 405 (1926). This paper is referred to later by *loc. cit.*

† See Schrödinger's papers in the *Ann. d. Phys.*, vols. 79—81 (1926). For the general method of obtaining matrices from the wave mechanics, see section 2 of the author's paper, *Roy. Soc. Proc.* A, vol. 112, p. 661 (1926).

Digitized by Google

where A is a dynamical variable, and the coefficients $A_{a'a}$ are functions of any steadily increasing variable such as t or t'. These coefficients are then the elements of the matrix that represents A. If we do choose just t' to be the variable of which the $A_{a'a}$'s are functions, then we get the matrix representation that is suitable for determining the radiation emitted in the direction (l_1, l_2, l_3). The matrix elements would now be of the form $ae^{2\pi i\omega t'}$, where the a's and ω's are constants, and we would then take the ω's to be the frequencies and the a's (if A is a component of total polarisation) to determine, according to classical formulae, the intensities of the components of radiation emitted in the direction (l_1, l_2, l_3).

§ 2. *Solution of the Wave Equation.*

We consider a free electron exposed to plane polarised incident radiation of wave number $\nu/2\pi$ moving in the direction of the x_1 axis with its electric vector in the direction of the x_2 axis. The Hamiltonian equation of the system is*

$$m^2c^2 = W^2/c^2 - p_1^2 - [p_2 + a' \cos \nu (ct - x_1)]^2 - p_3^2,$$

where p_1, p_2, p_3 and $-W$ are the variables canonically conjugate to x_1, x_2, x_3 and t, and a' is a constant, which may be assumed to be small. [p_2 differs from the ordinary component of momentum in the x_2 direction by the term $a' \cos \nu (ct - x_1)$.]

Schrödinger's wave equation is now

$$\{m^2c^2 - W^2/c^2 + p_1^2 + [p_2 + a' \cos \nu (ct - x_1)]^2 + p_3^2\}\, \psi = 0 \ldots (2),$$

in which the symbols p_1, p_2, p_3 and W mean the operators

$$-ih\frac{\partial}{\partial x_1}, \quad -ih\frac{\partial}{\partial x_2}, \quad -ih\frac{\partial}{\partial x_3} \text{ and } ih\frac{\partial}{\partial t}.$$

We must find the solutions ψ_a of this equation, and we can then determine the matrix representing a component X of the polarisation by obtaining an expansion for $X\psi_a$ of the form

$$X\psi_a = \Sigma_{a'} X_{a'a}\psi_{a'},$$

where the coefficients $X_{a'a}$ are functions of the single variable t'.

We now, as in the previous treatment of the problem, apply the linear canonical transformation

$$\left.\begin{aligned} x_1' &= ct - x_1 & p_1 &= -p_1' + l_1 W'/c \\ x_2' &= x_2 & p_2 &= p_2' + l_2 W'/c \\ x_3' &= x_3 & p_3 &= p_3' + l_3 W'/c \\ t' &= t - (l_1x_1 + l_2x_2 + l_3x_3)/c & W &= W' - cp_1' \end{aligned}\right\} \ldots (3).$$

The purpose of this transformation is two-fold; first it introduces explicitly the variable t' in terms of which we want the matrix

* *Loc. cit.*, equation (21).

Digitized by Google

elements, and secondly we shall be able to solve the wave equation in the new co-ordinates directly by separation of the variables. By substituting for p_1, p_2, p_3 and W in (2) their values given by the second set of equations (3), we find for the wave equation in the new co-ordinates, with neglect of terms involving a'^2,

$$\{m^2c^2 + 2W'/c \,.\, [(1-l_1)p_1' + l_2p_2' + l_3p_3' + l_2a' \cos \nu x_1'] + p_2'^2 + p_3'^2 + 2a'p_2' \cos \nu x_1'\} \psi = 0 \quad \ldots\ldots(4).$$

The symbols p_1', p_2', p_3' and W' here mean the operators

$$-ih\frac{\partial}{\partial x_1'}, \quad -ih\frac{\partial}{\partial x_2'}, \quad -ih\frac{\partial}{\partial x_3'} \text{ and } ih\frac{\partial}{\partial t'}.$$

Since the wave equation (4) does not involve x_2', x_3' or t' explicitly, it will have solutions of the form

$$\psi_\alpha = e^{i\alpha_2 x_2'/h} \,.\, e^{i\alpha_3 x_3'/h} \,.\, e^{-i\alpha_4 t'/h} \,.\, \chi(x_1'),$$

where α_2, α_3 and α_4 are arbitrary constants, and $\chi(x_1')$ is a function of the single variable x_1'. With this form for ψ_α we have

$$p_2'\psi_\alpha = \alpha_2\psi_\alpha, \quad p_3'\psi_\alpha = \alpha_3\psi_\alpha, \quad W'\psi_\alpha = \alpha_4\psi_\alpha,$$

and, more generally, if $f(p_2', p_3', W')$ is any function of p_2', p_3' and W', which may also involve x_1' and p_1' (f is thus an operator) we have

$$f(p_2', p_3', W')\psi_\alpha = f(\alpha_2, \alpha_3, \alpha_4)\psi_\alpha.$$

This rule applied to the left-hand side of equation (4) gives

$$\{2\alpha_4/c \,.\, [(1-l_1)p_1' + l_2\alpha_2 + l_3\alpha_3 + l_2a' \cos \nu x_1'] + b + 2a'\alpha_2 \cos \nu x_1'\} e^{i\alpha_2 x_2'/h} \,.\, e^{i\alpha_3 x_3'/h} \,.\, e^{-i\alpha_4 t'/h} \,.\, \chi(x_1') = 0,$$

where $$b = m^2c^2 + \alpha_2^2 + \alpha_3^2.$$

The factors $e^{i\alpha_2 x_2'/h}$, $e^{i\alpha_3 x_3'/h}$ and $e^{-i\alpha_4 t'/h}$ may be cancelled out, and a simple differential equation for the function $\chi(x_1')$ will then be left, whose solution is

$$\chi(x_1') = \exp -i\left\{\left(l_2\alpha_2 + l_3\alpha_3 + \frac{cb}{2\alpha_4}\right)x_1' + \frac{a'}{\nu}\left(l_2 + \frac{c\alpha_2}{\alpha_4}\right)\sin \nu x'\right\}\Big/(1-l_1)h.$$

The solution of equation (4) is thus

$$\psi_\alpha = \psi_{\alpha_2, \alpha_3, \alpha_4} = \exp i\left\{(1-l_1)(\alpha_2x_2' + \alpha_3x_3' - \alpha_4t') - \left(l_2\alpha_2 + l_3\alpha_3 + \frac{cb}{2\alpha_4}\right)x_1' - \frac{a'}{\nu}\left(l_2 + \frac{c\alpha_2}{\alpha_4}\right)\sin \nu x_1'\right\}\Big/(1-l_1)h \quad \ldots\ldots(5),$$

Digitized by Google

and since it contains the three arbitrary parameters α_2, α_3, α_4, and may also be multiplied by a fourth arbitrary parameter, it is a general solution.

§ 3. *The Integrals of the Equations of Motion.*

The first integrals of the equations of motion are obviously p_2', p_3' and W'. We must now determine the second integrals. We can do this in a way that is completely analogous to the way followed on the classical theory when one is solving a problem by the Hamilton-Jacobi method. Put

$$\psi_\alpha = \exp iS_\alpha/h \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(6).$$

Then S_α corresponds to the principal function in the Hamilton-Jacobi method, and $\partial S/\partial\alpha_r$ is a constant of integration of the system, where α_r is any one of the parameters occurring in ψ_α.

To prove this theorem on the quantum theory, differentiate (6) partially with respect to α_r. This gives

$$\frac{\partial\psi_\alpha}{\partial\alpha_r} = \frac{i}{h} e^{iS_\alpha/h} \frac{\partial S_\alpha}{\partial\alpha_r} = \frac{i}{h}\frac{\partial S_\alpha}{\partial\alpha_r}\psi_\alpha,$$

which shows that $\partial S_\alpha/\partial\alpha_r . \psi_\alpha$ is of the form

$$-ih\iiint c(\alpha', \alpha)\psi_{\alpha'} d\alpha_2' d\alpha_3' d\alpha_4',$$

where $c(\alpha', \alpha)$ is a certain very discontinuous "function" of α_2, α_3, α_4 and α_2', α_3', α_4', which is equal to zero except when $|\alpha_2 - \alpha_2'|$, $|\alpha_3 - \alpha_3'|$ and $|\alpha_4 - \alpha_4'|$ are very small, and is very large and positive when α_r' is a little greater than α_r, and very large and negative when α_r' is a little less than α_r. Now $-ihc(\alpha', \alpha)$ gives the elements of the matrix that represents $\partial S_\alpha/\partial\alpha_r$, and since these elements are all constants (*i.e.* independent of t'), $\partial S_\alpha/\partial\alpha_r$ must be a constant of integration of the system.

To apply this theorem in the present example, we must differentiate the index of the exponential in (5) partially with respect to α_2, α_3 and α_4. This gives

$$\left.\begin{aligned}
(1-l_1)x_2' - \left(l_2 + \frac{c\alpha_2}{\alpha_4}\right)x_1' - \frac{ca'}{\nu\alpha_4}\sin\nu x_1' &= c_2 \\
(1-l_1)x_3' - \left(l_3 + \frac{c\alpha_3}{\alpha_4}\right)x_1' &= c_3 \\
-(1-l_1)t' + \frac{cb}{2\alpha_4^2}x_1' + \frac{ca'\alpha_2}{\nu\alpha_4^2}\sin\nu x_1' &= c_4
\end{aligned}\right\} \quad \ldots\ldots(7),$$

where c_2, c_3 and c_4 are the second integrals. These formulae correspond to equations (35), (34) and (30) respectively in *loc. cit.*, which were there obtained by more laborious methods.

Digitized by Google

§ 4. *The Matrix Elements.*

We must now find the matrices that represent the components of the polarisation of the system. We need do this only for two components in directions that are perpendicular to the direction of emission (l_1, l_2, l_3) and mutually perpendicular. We can take these components to be the quantities

$$X = l_3 x_1 - \frac{l_2 l_3}{1-l_1} x_2 + \left(\frac{l_2^2}{1-l_1} - l_1\right) x_3$$

and

$$Y = l_2 x_1 + \left(\frac{l_3^2}{1-l_1} - l_1\right) x_2 - \frac{l_2 l_3}{1-l_1} x_3$$

multiplied by the electronic charge, as in the previous treatment of the problem*. Expressed in terms of the dashed variables, X and Y are given by

$$(1-l_1)\, X = l_3 ct' - l_3 x_1' + (1-l_1)\, x_3'$$
$$(1-l_1)\, Y = l_2 ct' - l_2 x_1' + (1-l_1)\, x_2'.$$

With the help of equations (7) we find

$$(1-l_1)\, X = c_3 + \frac{2\alpha_3\alpha_4}{b} c_4 + l_3 ct' + \frac{2(1-l_1)\alpha_3\alpha_4}{b} t' - \frac{2ca'\alpha_2\alpha_3}{\nu b \alpha_4} \sin \nu x_1'$$

$$(1-l_1)\, Y = c_2 + \frac{2\alpha_2\alpha_4}{b} c_4 + \frac{2(1-l_1)\alpha_2\alpha_4}{b} t' + \left(\frac{ca'}{\nu\alpha_4} - \frac{2ca'\alpha_2^2}{\nu b \alpha_4}\right) \sin \nu x_1' + l_2 ct' \qquad \text{.....(8).}$$

In these expressions for X and Y we know the matrices that represent all the terms except those involving $\sin \nu x_1'$. [The matrices representing the c's have already been discussed in the preceding section, while the matrix that represents t' is simply the diagonal matrix with diagonal elements t'.] Thus, in order to complete the matrix determination of X and Y, we have only to find the matrix that represents $\sin \nu x_1'$ and multiply it by the (constant) coefficients of $\sin \nu x_1'$ in the expressions (8). Since we are working only to the first order in a', and an a' occurs already in the coefficients of $\sin \nu x_1'$, we may neglect a' altogether when finding the matrix that represents $\sin \nu x_1'$.

With this approximation, we have from (5)

$$e^{i\nu x_1'} \psi_{\alpha_2, \alpha_3, \alpha_4} = \exp i \{(1-l_1)(\alpha_2 x_2' + \alpha_3 x_3' - \alpha_4 t') - [l_2\alpha_2 + l_3\alpha_3 + cb/2\alpha_4 + (1-l_1)h\nu]\, x_1'\}/(1-l_1)h$$
$$= \exp - i\nu' ct' . \exp i \{(1-l_1)(\alpha_2 x_2' + \alpha_3 x_3' - \alpha_4' t') - [l_2\alpha_2 + l_3\alpha_3 + cb/2\alpha_4']\, x_1'\}/(1-l_1)h,$$

* *Loc. cit.*, equations (33).

Digitized by Google

where ν' and α_4' are two new constants satisfying

$$\alpha_4 = hc\nu' + \alpha_4' \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (9)$$

and

$$\frac{cb}{2\alpha_4} + (1 - l_1)\, h\nu = \frac{cb}{2\alpha_4'} \quad \ldots\ldots\ldots\ldots\ldots\ldots (10).$$

We now have the relation

$$e^{i\nu x_1'} \psi_{\alpha_2, \alpha_3, \alpha_4} = e^{-i\nu' ct'} \psi_{\alpha_2, \alpha_3, \alpha_4'},$$

which gives an expansion for the left-hand side of the required type. It shows that the general element of the matrix representing $e^{i\nu x_1'}$, namely $e^{i\nu x_1'}(\alpha_2, \alpha_3, \alpha_4;\ \alpha_2', \alpha_3', \alpha_4')$, vanishes except when

$$\alpha_2' = \alpha_2, \quad \alpha_3' = \alpha_3 \quad \ldots\ldots\ldots\ldots\ldots\ldots (11),$$

and α_4 and α_4' satisfy the condition (10), and is equal to $e^{-i\nu' ct'}$ when these conditions are fulfilled. In the same way it may be shown that the general element of the matrix representing $e^{-i\nu x_1'}$, which we write for convenience $e^{-i\nu x_1'}(\alpha_2', \alpha_3', \alpha_4';\ \alpha_2, \alpha_3, \alpha_4)$, vanishes except when the same conditions (10) and (11) are fulfilled, in which case it equals $e^{i\nu' ct'}$. We can now write down without further trouble all the elements of the matrices representing X and Y.

§ 5. *Physical Interpretation of the Matrix Elements.*

The constant terms and the terms proportional to t' in the expressions (8) for X and Y will contribute nothing to the emitted radiation, and may be ignored. We are thus left with only the terms proportional to $\sin \nu x_1'$ to investigate. We see that their matrices contain only elements referring to transitions in which p_2', p_3' and W' change from a set of values α_2, α_3, α_4 to a set α_2', α_3', α_4' given in terms of α_2, α_3, α_4 by equations (10) and (11) (or from the α's to the α's), and further that the wave number of the radiation emitted during such a transition is $\nu'/2\pi$, where ν' is given by (9). Using the symbol Δ to denote the increase in a constant of integration of the system during a transition, we have from (11), (9) and (10)

$$\Delta p_2' = 0, \quad \Delta p_3' = 0, \quad \Delta W' = -hc\nu' \quad \ldots\ldots\ldots\ldots (12)$$

and

$$\Delta\,(cb/2W') = (1 - l_1)\, h\nu \quad \ldots\ldots\ldots\ldots (13).$$

These relations are sufficient to determine the recoil momentum of the electron and the frequency of the scattered radiation. Equation (13) or its equivalent equation (10) can be most conveniently used if we observe that, with neglect of a', p_1' is represented by a diagonal matrix whose diagonal elements α_1 are given

Digitized by Google

by $-ih$ times the coefficient of x_1' in the index of the exponential in (5), *i.e.*

$$\alpha_1 = -(l_2\alpha_2 + l_3\alpha_3 + cb/2\alpha_4)/(1-l_1)$$

so that
$$\Delta p_1' = \left(\frac{cb}{2\alpha_4} - \frac{cb}{2\alpha_4'}\right)\Big/(1-l_1) = -h\nu \quad \ldots\ldots\ldots(14).$$

The fact that we have to neglect a' in order to be able to give a meaning to $\Delta p_1'$ corresponds physically to the fact that in order to give a meaning to the recoil momentum of the electron we must neglect the oscillations of the electron due to the incident radiation. Equations (12) and (14) give, when written in terms of the undashed variables,

$$\Delta p_1 = h\nu - l_1 h\nu'$$
$$\Delta p_2 = -l_2 h\nu'$$
$$\Delta p_3 = -l_3 h\nu'$$
$$\Delta W/c = h\nu - h\nu',$$

which are just the equations that express the conservation of momentum and energy on Compton's light-quantum theory of scattering. (We are neglecting a' again when we count p_2 as an ordinary momentum.)

The elements of the matrix that represents the periodically varying part of X are

$$X(\alpha_2, \alpha_3, \alpha_4;\ \alpha_2', \alpha_3', \alpha_4') = \frac{ica'\alpha_2\alpha_3}{\nu(m^2c^2 + \alpha_2^2 + \alpha_3^2)\alpha_4(1-l_1)}\, e^{-i\nu' ct'}$$

and

$$X(\alpha_2', \alpha_3', \alpha_4';\ \alpha_2, \alpha_3, \alpha_4) = \frac{-ica'\alpha_2'\alpha_3'}{\nu(m^2c^2 + \alpha_2'^2 + \alpha_3'^2)\alpha_4'(1-l_1)}\, e^{i\nu' ct'},$$

where the α's are given in terms of the α's by (10) and (11). The product of these two matrix elements gives a quarter of the square of the amplitude, which, by insertion in the formula of the classical theory, determines the intensity of the emitted radiation. The amplitude squared is thus

$$C_X^2 = 4\,\frac{c^2a'^2\alpha_2^2\alpha_3^2}{\nu^2(m^2c^2 + \alpha_2^2 + \alpha_3^2)^2\alpha_4\alpha_4'(1-l_1)^2} \quad \ldots\ldots\ldots(15).$$

Instead of actually evaluating the intensity from this formula, we can obtain results more simply by comparing C_X^2 with its value according to the classical theory. This classical value is given by one putting $h = 0$ in (15), which comes to writing α_4 for the α_4' in the denominator. Hence the quantum value of C_X^2 is just α_4/α_4' times its classical value. This result may easily be verified to hold also for C_Y^2, the amplitude squared of Y.

Digitized by Google

From (9) and (10) we have

$$a_4 - a_4' = hcv'$$

and

$$\frac{1}{a_4'} - \frac{1}{a_4} = \frac{2(1-l_1)}{cb} h\nu,$$

which give, by division,

$$a_4 a_4' = \frac{c^2 b}{2(1-l_1)} \frac{\nu'}{\nu} \quad \text{.....(16).}$$

If the electron is initially at rest, we have initially

$$p_1 = p_2 = p_3 = 0, \quad W = mc^2,$$

from which, using the transformation equations (3), we find for the initial values of p_2', p_3' and W',

$$a_2 = -\frac{l_2 mc}{1-l_1}, \quad a_3 = -\frac{l_3 mc}{1-l_1}, \quad a_4 = \frac{mc^2}{1-l_1}.$$

This gives

$$b = m^2c^2 + a_2^2 + a_3^2 = \frac{2m^2c^2}{1-l_1},$$

so that from (16)

$$\frac{a_4}{a_4'} = \frac{2(1-l_1)}{c^2 b} \frac{a_4^2}{a_4'} \frac{\nu}{\nu'} = \frac{\nu}{\nu'}.$$

The amplitude squared is thus greater than its value according to the classical theory in the ratio ν/ν'. Since the intensity is proportional to the amplitude squared multiplied by the fourth power of the frequency, and the frequency of the scattered radiation is less than its value on the classical theory in the ratio ν'/ν, it follows that the intensity is less than its value on the classical theory in the ratio $(\nu'/\nu)^3$. This result is the same as that obtained in *loc. cit.*, and is in good agreement with experiment*.

* *Note added in proof*:—A paper by W. Gordon dealing with the same subject has recently appeared [*Zeits. f. Phys.*, vol. 40, p. 117 (1926)]. Gordon's method makes use of an expression for the electric density for determining the field produced by the scattering electron, and differs from the method of the present paper, in which the wave equation is used merely as a mathematical help for the calculation of the matrix elements, which are then interpreted in accordance with the assumptions of matrix mechanics.

Digitized by Google

On the polarisation of mercury lines emitted from a discharge tube in a magnetic field. By Dr H. W. B. SKINNER*, Trinity College.

[*Received* 17 August, *read* 22 November 1926.]

§ 1. *Introduction.*

Theoretical work on the Zeeman effect has led to the conclusion that the light emitted by atoms in a magnetic field should be unpolarised; that is that the intensity of the Zeeman components polarised with the maximum electric vector parallel to the field will be equal to the intensity of those polarised perpendicular to the field. This would be the case on the classical theory; on the quantum theory it can be shown to follow from the principle of spectroscopic stability†. It may be shown that the equality of the intensities of the components polarised parallel and perpendicular to the field is related to the absence of double refraction in a gas whose molecules are orientated by the application of a magnetic field, and this last property has been established by Gerlach and Shütz‡.

On the other hand, some old experiments by Egoroff and Georgiewski § and by Wiechern ‖ have shown that the light emitted from salted flames and discharge tubes in a magnetic field of about a thousand gauss is partially polarised with the direction of the maximum electric vector perpendicular to the line of force, the polarisation varying in a complex way with the field strength. It seemed therefore probable that the observed polarisation in these cases might be due to the fact that the source of light was not an ideally simple one. In fact, that the light actually observed was not that directly emitted from atoms subject to isotropic conditions of excitation. It was with a view to clearing up this question that the present experiments were performed.

§ 2. *Experimental Methods.*

To construct an ideally simple source of light (in the meaning of the above definition) is not an easy matter. If one uses a discharge tube, one introduces a certain directedness into the conditions for the excitation of the atoms; if a flame, one knows that the conditions on the outside of the flame are different from those within it.

* Exhibition of 1851, Senior Student.
† Heisenberg, *Zs. f. Phys.* vol. xxxi.
‡ *Naturwiss.* vol. xi. p. 617 (1925).
§ *C.R.* vol. 124, pp. 748, 749; vol. 125, p. 16 (1897).
‖ *Göttingen Diss.* (1914).

Digitized by Google

The chief requisite for the present experiment is to have a source of light which, if not ideally simple, can at any rate be controlled. With this object the source chosen was a low-pressure electron-maintained arc in mercury vapour. The advantage of this source is that the low pressure may be expected to remove a number of disturbing conditions, and that the current passing in the tube can be varied.

Fig. 1 shows ($\frac{1}{5}$th actual size) the construction of the arc which runs on mercury vapour in equilibrium at atmospheric temperature, the pressure being therefore about one-thousandth of a mm.

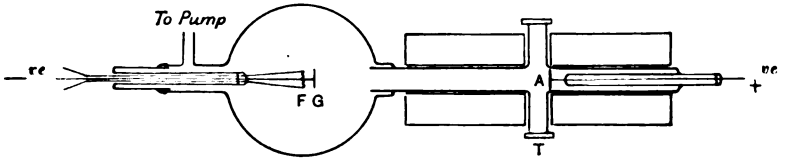


Fig. 1.

A potential difference of 110 volts is applied between the tungsten filament F and the anode A, the grid G being only used for starting the arc. A current of up to $\frac{1}{2}$ amp. can be passed between F and A. The arc passes through the bored poles of a small electromagnet which can provide a field of up to 3000 gauss in the space between the pole pieces which were about 17 mm. apart. The light is observed through the cross tube T which passes between the poles.

The tube was continuously exhausted by a rotating Garde pump which also supplied the mercury vapour.

The discharge is of a blue-white colour and in the absence of a magnetic field fills the tube with a uniform glow. The application of the field causes the concentration of the glow into a line (whose diameter depends on the field strength) passing down the centre of the tube and the light emitted from points outside this line is relatively weak.

The light from the arc was analysed by a set of mercury lamp filters, which were, on the whole, more convenient to use than a spectroscope. But when greater resolution was necessary, a spectroscope was used.

For the detection of polarised light, a Savart plate with analysing Nicol prism was used. When polarised light falls on it a set of parallel fringes is seen. By a slight motion of the plate the fringes may be caused to move and much fainter fringes may be detected in this way than if they are stationary.

Digitized by Google

The measurement of the polarisation was carried out by compensating the polarisation by causing the light to pass through a glass plate capable of being rotated about a suitably orientated axis. From the angle at which the glass plate must be placed to neutralise the fringes, the percentage polarisation of the light emitted can be calculated. The plane of polarisation is found by rotating the Savart plate (with the glass plate away) until the fringes are seen with maximum strength. This does not determine the plane of polarisation uniquely, but tells one, for example, that it is either horizontal or vertical. By the use of the glass plate it is easy to find out which of these alternatives is correct.

The results will be expressed as "percentage of polarisation." By this is meant the quantity

$$\Pi = 100\,\frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + I_{\perp}},$$

where $I_{\parallel}$, $I_{\perp}$ are the intensities of the light emitted as vibrations with the electric vector parallel or perpendicular to a given direction, the plane of polarisation. With the experimental arrangements, and suitable source of light such as the green or yellow mercury lines, a polarisation of $\frac{1}{2}\,\%$ could be detected. In the case of violet lines, however, the accuracy is, owing to the poor visibility, considerably less.

§ 3. *Results of Experiments.*

The lines which have been investigated are the mercury lines 4358 ($2\,^3P_1 - 2\,^3S_1$), 5461 ($2\,^3P_2 - 2\,^3S_1$), 5770 ($2\,^1P_1 - 3\,^3D_2$) and λ5791 ($2\,^1P_1 - 2\,^1D_2$). With the exception of the first, these lines all show polarisation effects which vary with the field strength. We shall consider the lines separately and the results, except where otherwise stated, will refer to observation in a direction perpendicular to the lines of magnetic force.

(*a*) λ 4358.

Observations on this line are rather difficult owing to poor visibility. There is also a certain amount of unavoidable contamination with the fainter line λ 4347 ($2\,^1P_1 - 4\,^1D_2$). However, under no conditions was any polarisation observed on λ4358, though $2\,\%$ would have been detected.

(*b*) λ 5461.

This line is unpolarised in a zero field, the accuracy of detection being about $\frac{1}{2}\,\%$ of polarisation. In a stronger field, however, the light becomes polarised perpendicular to the lines of force, the polarisation increasing with increasing field strength

Digitized by Google

until a maximum is reached in a field of 2000 Gauss. A curve showing the effect is given in Fig. 2.

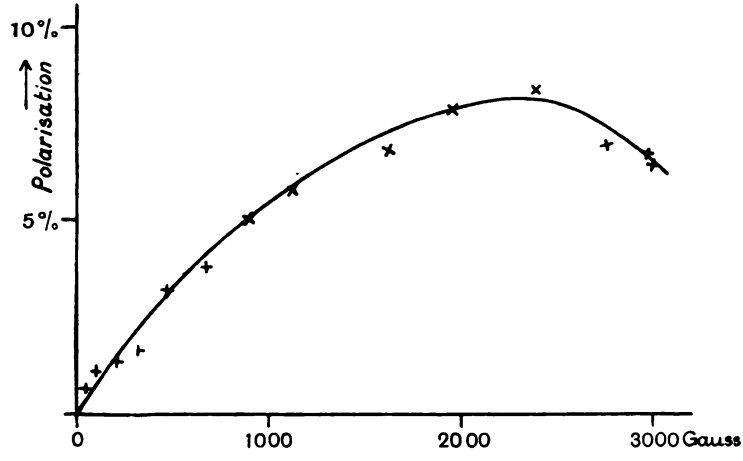


Fig. 2.

The polarisation however also depends on other factors than the magnetic field. In Fig. 2 the current in the tube was kept

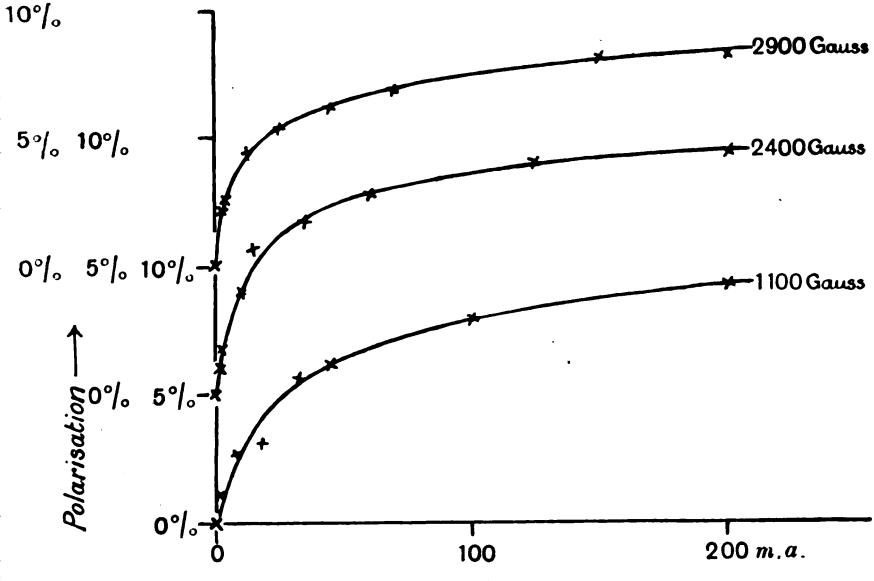


Fig. 3.

Digitized by Google

constant. If, instead, we vary the current, keeping the field constant, we obtain the curves shown in Fig. 3.

These facts show that the effects depend on the magnitude and distribution of the current over the cross section of the tube. It has been mentioned that the magnetic field exerts a concentrating action on the discharge, condensing it into a line down the centre of the tube. The precise diameter of this line depends on the field and also on the current through the tube, and the concentration does not necessarily proceed continuously with increasing magnetic field. The exact character of the concentration seems to depend on relatively small factors in the design of the tube. Consistent values of polarisation may be obtained if the tube is unaltered, but after a slight alteration the magnitude of the polarisation may change, though the general character of the curves remains the same. To take a definite example, a maximum value of 6 °/₀ instead of the 9 °/₀ of Fig. 3 was obtained in one instance. For this reason, the curves given must be taken simply as typical. An important general fact is however that the polarisation falls to zero as the current in the tube is diminished, whatever the value of the magnetic field.

A further fact which has been established is that the polarisation increases when the angle between the field and the direction of observation is reduced. At an angle of 72° to the lines of force the polarisation found was 6·3 °/₀, the corresponding value in a direction perpendicular to the lines of force being 3·6 °/₀. This seems to be due to the fact that the thickness of the discharge through which the light passes on its way out of the tube, is greater when observed at an acute angle.

If a small quantity of air is allowed to enter the tube the polarisation is reduced. With a pressure of the order of $\frac{1}{100}$th mm. the reduction under standard conditions of current and magnetic field was found to be from 6 °/₀ to 1 °/₀. The observations in this case were spectroscopic since it is necessary to separate the light of the green line from that of neighbouring bands. Similar results were also obtained with H_2. The significance of this result will be discussed subsequently.

(*c*) λ 5770, λ 5791.

In the case of these lines, no polarisation was observed in fields of 200 gauss or more. On the other hand, in a zero or weak field, a polarisation is observed, the plane of polarisation (electric vector) being in the direction of the discharge. A polarisation of 7 °/₀ was observed in a zero field, and this amount was approximately independent of the current in the tube.

It is evident that this effect is of quite a different character from that found in the case of the green line. It forms the subject

Digitized by Google

of a further paper* in which it is proved to be due to the directed character of the discharge which causes the excitation of a line to be imperfectly isotropic. It is shown that in the case of the excitation of these lines by a perfectly directed stream of electrons, the polarisation may reach a much higher value. In the type of arc used in the present experiments, electrons are travelling in all directions, but the applied potential serves to introduce a certain general drift on to the motions of these electrons and to this cause the observed polarisation is due.

§ 4. *Discussion of Results.*

The experiments have therefore given evidence of two distinct polarisation effects. One of these, the effect found for the yellow lines, is destroyed by the application of a sufficiently strong magnetic field, and in this effect we are concerned with a case of anisotropic excitation; for it depends on the directed character of the discharge. The other effect, that on the green line, is only found in a sufficiently strong magnetic field, and is a complex effect whose magnitude is determined by the precise conditions of the discharge, and in particular vanishes to zero when the current in the tube is reduced.

On the other hand, in the case of the yellow and violet lines, no polarisation at all which is due to the magnetic field was observed.

We may thus state that the experiments give definite evidence that the intensity relation of the Zeeman effect mentioned in § 1, namely that the intensity of the components polarised parallel to the field is equal to the intensity of the components polarised perpendicular to the field would be found true if sufficiently simple conditions of observation could be obtained. To establish this result was the main object of the present work.

It remains now to give a suggestion as to the cause of the polarisation observed in the case of the green line. It may be stated that the polarisation-field-strength curve given in Fig. 2 is of the same type as those obtained by Wiechern.

We have seen that the polarisation appears to be caused by the conditions in the gas through which the light passes on its way out of the tube. There are various reasons for supposing absorption of the light by atoms in the outer parts of the discharge to play a part. In the first place, the effect has only been observed in the case of the green line $\lambda 5461$ ($2\,^3P_2 - 2\,^3S_1$). Now the $2\,^3P_2$ state is a metastable state, and an atom in this state can only pass out of it by a collision. At the low pressures concerned there will be a considerable concentration of atoms in this state ready to absorb $\lambda 5461$. On the other hand, in the case of

* *Proc. Roy. Soc.* 112A, p. 642 (1926).

Digitized by Google

the remaining lines investigated, no metastable states are concerned. Again, the introduction of a foreign gas will reduce the number of atoms in the $2\,^3P_2$ state and, as we have seen, it also reduces the polarisation. This hypothesis will also explain the fall of the polarisation to zero as the current diminishes, for with small currents the number of atoms in the $2\,^3P_2$ state will be small.

Absorption will cause a self-reversal of the line, and was long ago suggested by Cotton* to explain the results of Egoroff and Georgiewski mentioned in § 1.

These authors also found that the lines most easily reversed show the greatest polarisation effects.

To account for the production of polarised light by absorption, one must suppose that the conditions under which the light is emitted are different from those in the layer where it is absorbed. Now, owing to the concentration of the discharge by the magnetic field, most of the light is emitted from the central part of the tube, while the absorption presumably takes place indiscriminately anywhere in the tube.

The most likely supposition to account for a difference in the conditions under which emission and absorption take place would seem to be that it is due to a very slight non uniformity of the magnetic field (such as must exist on account of the hole bored in the pole pieces). A hypothesis of this kind was also put forward by Cotton. Such a non-uniformity will imply that the wave-lengths of the Zeeman components of the directly emitted light will not exactly correspond with the wave-lengths which would be most strongly absorbed by the absorbing atoms. This divergence will be more marked for the outer Zeeman components (which are polarised perpendicular to the field) than for the central components (polarised parallel to the field). Thus the light emitted polarised parallel to the field will be more absorbed than that emitted polarised perpendicular to the field, and the total light emitted will show traces of polarisation in a direction perpendicular to the field. This is what is observed, and it would appear that this hypothesis is sufficient to account for the general characteristics of the effect.

Summary.

1. Polarisation measurements are described on the Hg lines λ 4358, 5461, 5770, 5791 emitted from a low-pressure electron-maintained arc to which a longitudinal field of up to 3000 gauss may be applied.

* *C.R.* 125, p. 1169 (1897).

Digitized by Google

2. Two types of polarisation effect have been found. The first is typified by the line λ 5461. It is a polarisation of the light with the electric vector perpendicular to the lines of force. Its maximum value is 9 %, but it decreases to zero (*a*) with decreasing field strength and (*b*) with decreasing current in the tube. It is suggested that this effect is due to a self-reversal of λ 5461 owing to the metastability of the $2\,^3P_2$ state and that the polarisation may be caused through a slight non-uniformity of the magnetic field.

The second type of effect is found for λ 5770 and 5791. The polarisation is in a plane parallel to the discharge, and is found in a zero magnetic field. This effect is further discussed elsewhere.

The line λ 4358 is unpolarised under all conditions tried.

3. Neither of these effects implies a polarisation of the total light emitted directly from atoms in the Transverse Zeeman Effect when the atoms are excited under isotropic conditions and the experiments give definite evidence against the existence of such a polarisation effect.

In conclusion, I should like to express my thanks to Professor Sir Ernest Rutherford, O.M., P.R.S., for his continual interest in the work.

Digitized by Google

Note on the Effect of Alpha Particles on Paraffin. By WILLIAM T. RICHARDS, Ph.D. (Communicated by Prof. Sir E. RUTHERFORD.)

[*Received* 25 August, *read* 25 October 1926.]

Duane and Scheuer (*Le Radium*, **10**, 33, 1913) found, by measurements of the amount of gas mixture evolved by the action of alpha particles on water and ice, that the number of molecules acted upon divided by the number of ion pairs formed was 0·96 in the former case and only 0·05 in the latter. Since this seemingly anomalous result is the only satisfactory published comparison between the behaviour of solids and liquids under the action of alpha particles, a short research of the same type has been undertaken with a substance of somewhat different properties. Paraffin was chosen for this purpose because of the vigorous action to be expected (cf. Lind, *Science*, **40**, 304 (1924) footnote), because of its simple empirical formula, because a previous investigation of so complex a molecular structure had not been undertaken, and because it may be obtained liquid and solid at ordinary temperatures without important change in chemical composition.

Apparatus.

The apparatus consisted, in essence, of a reaction chamber connected by a short capillary tube to a standardized burette which, in turn, was connected to a gas pipette and eudiometer chain terminating in a Ramsay pipette. The eudiometer served also as a means of evacuation. Control measurements showed that quantities of gas of the order of 1 c.c. could be measured with an error of about 0·004 c.c. Since measurements were made over water to secure constant aqueous tension in the burette system, any water vapour evolved by the alpha particle action could not have been detected

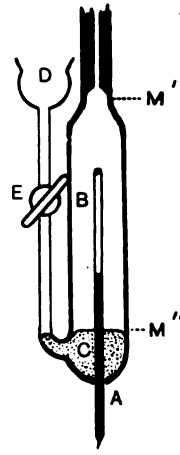


Fig. 1.

The reaction chamber is illustrated in Figure 1. It was made of glass, having a capacity between the limits M' and M'' of about 1·0 c.c. A represents an orifice which permitted the introduction of an alpha ray tube B, and was subsequently tightly sealed with hard wax. The wax was protected from the action of alpha rays by a layer of mercury C which was introduced through the funnel tube D. When liquid paraffin was used it was introduced through D to the level M', and the tap E closed.

Digitized by Google

When solid paraffin was used it was melted on the inner wall of the reaction chamber in a uniform coating before the tube B was in place, and the side tube leading to D was then filled with mercury.

The liquid paraffin was obtained commercially as "refined medicinal paraffin." The solid was a good grade of hard paraffin wax with a melting point at about 50° C. The two are not, of course, chemically identical, but it is necessary only that they should both have the empirical formula C_mH_n, with m and n of comparable magnitude in each case, for the results to have significance. Impurities to the amount of 5 % would not have affected the results.

Experiments.

Solid Paraffin. 48·3 mg. of emanation were employed to cause the action, contained in an α-ray tube having a range of 4·5 cm. $\pm$ 0·5 cm. The colour of the substance changed, after long continued exposure to the radiation, from white to a brown tinged with violet.

Table I, in which θ represents time in hours and v volume in c.c. of gas evolved, shows the measurements obtained calculated (with appropriate temperature and aqueous tension corrections) for 760 mm. and 12·6° C.

TABLE I.

θ	v	$\frac{\Delta v}{\Delta \theta}$	$\log_{10} \frac{\Delta v}{\Delta \theta}$
0	0	—	—
19·5	0·58	0·0298	$\bar{2}$·474
91·3	2·04	·0203	$\bar{2}$·308
114·8	2·24	·0085	$\bar{3}$·930
140·5	2·42	·0070	$\bar{3}$·845
163·8	2·55	·0056	$\bar{3}$·747

These measurements are represented graphically in Figure 2. The second point has been omitted since it involves too long an interval of time to give a reliable average value.

A series of measurements was also made on the pressure change at constant volume due to emanation in a small evacuated tube lined with paraffin. This method has the advantages of eliminating the uncertainty of range of the α-ray tube, and of showing the additional effect due to the action of recoil atoms. Since, however, the results obtained show no tendency not revealed in Table I, and since corresponding measurements cannot be made with liquids, they will not be detailed. 0·35 c.c. of gas were evolved per mg. of emanation decayed over a period of 29 hours, the volume being given under standard conditions.

Digitized by Google

Liquid Paraffin. 22·0 mg. of emanation were employed in an α-ray tube having a range of $4{\cdot}5 \pm 0{\cdot}5$ cm. (unfortunately not the same tube as that used in the solid determination). Fifteen minutes after the introduction of the tube small bubbles, which had formed on its surface, began to rise slowly through the liquid. A microscopic examination of the surface of the tube failed to show the action of single alpha particles. After two weeks a black shell, resembling carbon in appearance and properties, collected on a small part of the tube surface. Table II shows the measurements obtained, v being expressed at 12·6° C. and 760 mm. as before.

TABLE II.

θ	v	$\frac{\Delta v}{\Delta \theta}$	$\log_{10} \frac{\Delta v}{\Delta \theta}$
0	0	—	—
1·42	0·033	0·024	$\bar{2}{\cdot}380$
5·33	·080	·012	$\bar{2}{\cdot}079$
23·17	·139	·0033	$\bar{3}{\cdot}519$
26·67	·151	·0033	$\bar{3}{\cdot}519$
48·00	·188	·0018	$\bar{3}{\cdot}255$
100·00	·225	·0007	$\bar{4}{\cdot}851$
144·50	·249	·0006	$\bar{4}{\cdot}740$

These measurements also are represented graphically in Figure 2.

Errors. That the deviation from linearity of the rate of evolution of the gas in liquid paraffin is not due to systematic error in the apparatus is apparent by a careful consideration of the method of procedure, and is confirmed by the much greater regularity in the solid, where the method was identical. That it is not due to the solubility of hydrogen in paraffin is obvious, as this would influence the results in the opposite direction; the invalidity of this objection is confirmed by the immediate appearance of bubbles which rose very slowly, undiminished in size, to the surface of the liquid. The greatest error in the quantitative comparison of these results is the ranges of the two α-ray tubes which could be determined, as usual, within 10 °/₀ only. This does not, however, affect the rate of evolution of gas in each determination.

Analysis of the gas. Attempts to analyse the exceedingly limited quantity of gas obtained indicated that it was only hydrogen. A satisfactory analysis was not obtained, however, and the nature of the gas is of no importance for the main thesis of the work. It must be borne in mind here, as in other parts of this paper, that radium emanation can be obtained only in limited quantities, and that experiments with it cannot be repeated with the ease and precision possible in most chemical processes.

Digitized by Google

Comparison of Results.

The following results are the basis for comparison:

State	θ	$(mg)_0$	Δmg	v	$\frac{v}{\Delta mg}$
Solid	163·8	48·3	34·2	2·440	0·0713
Liquid	144·5	22·0	14·6	0·238	0·0163

Here θ represents the number of hours over which action was observed, $(mg)_0$ the number of mg. of radon at the commencement of action, Δmg the amount of radon decayed during action, and v the volume of gas generated expressed under standard conditions.

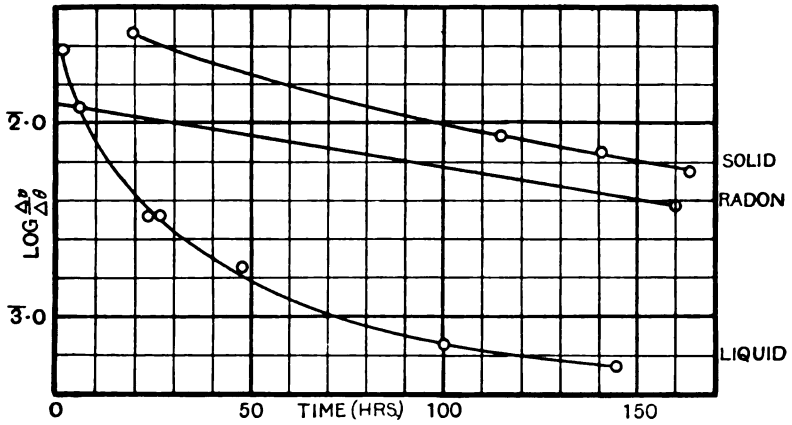


Fig. 2.

In the discussion below, the action due to β and γ rays is assumed to be negligible.

The values for $\frac{v}{\Delta mg}$ given above show, even as they stand, a striking difference from the results of Duane and Scheuer (*loc. cit.*) on ice and water. Whereas these investigators found only $\frac{1}{20}$ of the action in the solid state which they observed in the liquid, the variation here observed is by a factor of 4 in the opposite direction. A part of this latter divergence is clearly attributable to the mutual irregularity in time-action coefficient made clear by a comparison of the curves in Figure 2. This remains to be explained.

In order to treat this irregularity it is necessary to make only the almost axiomatic assumption that chemical action observed is proportional not only to the amount of emanation decayed (Ramsay's Law), but also to the amount of substance present if the latter be

Digitized by Google

small. From this, and from the following measurements, may be calculated the relative numbers of molecules instantaneously accessible to α-rays in solid and liquid. The measurements are: inside diameter of the reaction chamber, 0·62 cm.; its length, 3 cm.; length of active portion of the α-ray tube, 1·5 cm.; its diameter 0·1 cm. Taking the range of the average alpha particle in paraffin, calculated from Bragg's Law, to be ·0024 cm., the ratio $\frac{0{\cdot}0141 \text{ c.c.}}{0{\cdot}0012 \text{ c.c.}}$ is obtained. Thus 12 times as many molecules of solid as of liquid paraffin are subjected to $\frac{34{\cdot}2}{14{\cdot}6}$ or 2·3 times as many mg. of emanation, and we should expect to find $\frac{12}{2{\cdot}3}$ or 5·1 as great a volume of gas generated from the solid as from the liquid under the conditions of the experiment. This number is in fair agreement with the $\frac{7{\cdot}13}{1{\cdot}63}$ or 4·4 ratio actually observed for the gas volume generated in the two cases. The calculated value is based on the further assumption that mixing is so incomplete as to be negligible in the liquid. This is partly justified by the concordance obtained; it is further substantiated by the high viscosity of the liquid, and by a consideration of the nature of its decomposition products.

Thus, with a liquid of high viscosity, there appears to be an accumulation of decomposition products in the volume of alpha particle action which acts as a partial screen without producing gas in doing so. If this is virtually unaffected by such weak mixing agencies as are available in the liquid, it must be more noticeable at the surface of the tube than at the outer circumference of the volume of action, and the ultimate black coating of a part of the tube is accounted for. A modification of the usual procedure for the decomposition of liquids for very viscous liquids is therefore seen to be necessary. The slight deviation from linearity of solid paraffin may be even more plausibly accounted for on the basis given above.

Corrections for this effect in the liquid may be made from data already at hand. An extrapolation along the line of emanation decay from the initial rate of gas evolution should give the rate of evolution at the discontinuation of the experiment if the supply of liquid paraffin had been unlimited. If the initial rate is taken as that observed between 1·42 and 5·33 hours, when lack of adjustment to temperature and aqueous tension changes has surely been overcome, the result fixes a minimum for the quantity of gas so evolved. Such an extrapolation, which is indicated in Figure 2, gives a value of 1·08 c.c. at 12·6° C. for 160 hours of action by 15·4 mg. of emanation, or 0·0674 c.c. per mg. emanation at 0° C.

Digitized by Google

On the other hand, a correction may be made on the basis of the number of molecules available to action, as calculated above, which gives $5{\cdot}1 \times 0{\cdot}0163$ or 0·0830 c.c. of gas per mg. emanation. Although the methods by which each of these corrected values has been obtained are open to question, their average result, 0·0702 c.c. per mg. emanation, agrees well with the 0·0712 c.c. which was found in the solid determination, and must be given more weight than the uncorrected value. Evidently, then, the state of aggregation does not play an important part in the chemical effect of alpha particles on some substances, and divergencies from this conclusion must be attributed to secondary action wherever possible. This carries with it the corollary that the ionization relations in solids and liquids are comparable, a result which has been demonstrated for gases and liquids by previous investigations of a similar character.

A quantitative reconciliation between these results and those of the previous investigators on water cannot here be attempted. The divergence may, however, be descriptively accounted for by a consideration of the chemical and mechanical aspects of the two reactions studied. The mobility in water is sufficiently great to effectively separate a large proportion of the decomposition products which, in ice, remain practically without change of position. Whether it be supposed that they combine according to the laws of probability and their space distribution after formation, or whether recombination into their respective molecules be considered from the point of view of the free energies of the reactions involved, or both, the reformation of water is clearly favoured over paraffin under approximately similar conditions. Hence, with the latter, the mobility of the medium is not important.

According to Geiger (*Proc. Roy. Soc.* **83** A, 565 (1910)), a diminution of the paths of the alpha particles from the decay of radon by 2·6 cm. of air should affect their combined energy by a factor of 0·6, whereas the experiments with solid paraffin at constant volume indicate that 0·2 of the destructive energy of radon is effective with an α-ray tube of 4·5 cm. range. This discrepancy must be laid to the fact that the common method of measuring range gives the maximum value only; that this does not correspond to the average range is demonstrated by the collection of black coating on only a small part of the tube surface after long action in liquid paraffin. This is an added argument against the quantitative measurement of chemical action by α-ray tubes; they may be made to give comparable values, but their absolute values are uncertain. It is interesting to note that the amount of gas, 0·35 c.c. per mg. emanation, obtained from solid paraffin without an α-ray tube, corresponds closely to the value 390 c.c. per curie found by Duane and Scheuer (*loc. cit.*) in the decomposition of water, since it seems

Digitized by Google

to indicate a general similarity in the ionic mechanism of the two processes.

Lecture Demonstration.

A small trough of two microscopic cover glasses enclosing, between suitable cemented metal supports, an α-ray tube containing 20 or more mg. of emanation, shows clearly on the screen rising bubbles due to the intense chemical action when filled with liquid paraffin and mounted in a projecting lantern focussed on the tube.

Summary.

1. It has been shown that the action of alpha particles on paraffin is independent of the state of aggregation of the substance.

2. Defects in the customary procedure for the decomposition of liquids and solids by the emanation tube method are discussed.

3. A lecture demonstration of the chemical effect of alpha particles is described.

It is a pleasure to express my thanks to Sir Ernest Rutherford for calling my attention to the violent action of alpha particles on paraffin, and to the International Education Board whose Fellowship made possible my short stay in the Cavendish Laboratory.

Digitized by Google

The application of the method of the magnetic spectrum to the study of secondary electronic emission. By C. F. SHARMAN, B.A., King's College. (Communicated by Prof. Sir E. RUTHERFORD.)

[*Received* 25 August, *read* 22 November 1926.]

There are two ways of investigating the velocity distribution of an electronic emission—the retarding potential and the magnetic deflection methods. These have been applied to the case of the photo-electrons by Millikan* and Ramsauer† respectively. On one important point the results disagreed; Millikan found a definite maximum velocity, while Ramsauer obtained an asymptotic falling off of the number of electrons with increasing velocities. In a critical discussion of the two methods Klemperer‡ has shown that they agree when photoelectric activity of the collecting electrode in the first, and when electronic reflection at the walls and slits in the second, are eliminated.

The nature of the results yielded by the two methods shows the essential difference. The use of retarding potentials gives the number of electrons with energies greater than any given value, whereas the magnetic spectrum gives the number between limits of velocity the extent of which depends only on the geometry of the apparatus. The latter is, in fact, the differentiated form of the former, and is for obvious reasons the more delicate means of investigating the distribution problem. For instance, in the magnetic spectrum a homogeneous group of electrons of small intensity appears as a definite peak, whilst the other method gives a slight change of slope in the voltage-current curve which would probably be smoothed out as an experimental fluctuation. Moreover, the transformation of this voltage-current curve to its differentiated form is very inaccurate.

Ramsauer used the ordinary semi-circular focussing method, developed for beta-ray spectra, but substituted a Faraday cylinder for the photographic plate; he measured the current to the cylinder for various values of the deflecting magnetic field. Owing to the smallness of these currents he used very wide and long slits and corrected the distribution so obtained to that which would have been found with very small slits. The corrections were surprisingly small. The drawback of large apertures is that the diffuse rays which have not travelled in circular paths but which reach the Faraday cylinder are greatly increased; these destroy

* Millikan, *Phys. Rev.* vol. VII. p. 356 (1916).
† Ramsauer, *Ann. der Phys.* vol. XLV. p. 961 (1914).
‡ Klemperer, *Zeits. für Phys.* vol. XVI. p. 280 (1923).

Digitized by Google

the definition of the homogeneous groups and produce a continuous background.

Recently, Becker has applied the retarding potential method to the investigation of the velocity distribution of the secondary electrons excited by alpha-rays* and by a primary electron beam†. He found the two distributions to be identical and claims that they are both Maxwellian. The present paper deals with the application of the method of the magnetic spectrum to the case of the secondaries excited by a primary electron beam. The following are the points of chief importance: (1) The production of vacua sufficiently good that no distorsion of the spectrum occurs owing to the absorption of the slower electrons by the residual gas. (2) The reduction to a minimum of electron reflection from the sides of slits and the walls of the container. (3) The elimination of stray magnetic fields and the accurate control of the applied field, since the value of the field is an important factor in the form of the final energy distribution curve.

Apparatus.

The primary electron beam was produced by accelerating the thermions from a helical Tungsten filament F (Fig. 1) by charging it to a negative potential varying between 100 and 1000 volts.

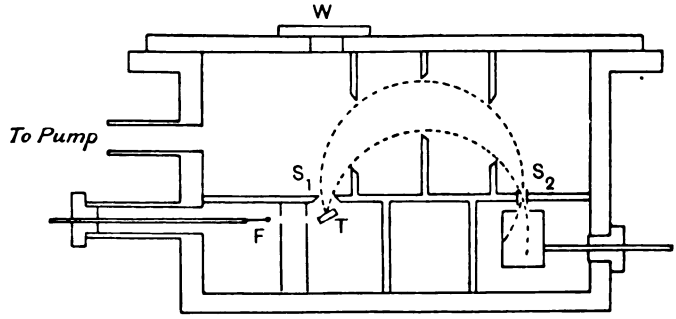


Fig. 1.

This beam passed through two earthed slits (1 mm. × 4 mm.) and fell on the target T, which was supported so as to make an angle of 45° with the beam and so that the part exposed to the primary rays was immediately below the first defining slit S_1. The adjustment of the target for position was checked by viewing the hot filament by reflection in the target through the slit S_1 and a small plate glass window W. In the majority of experiments the

* Becker, *Ann. der Phys.* vol. LXXV. p. 217 (1924).
† Becker, *Ann. der Phys.* vol. LXXVIII. p. 253 (1925).

Digitized by Google

target was a copper plate with a polished surface, which was earthed by connecting to the inside of the containing box. Of the total secondary emission from the target, the part passing through the slit S_1 was used for the determination of the distribution. The second defining slit S_2 was made with deep sides parallel to the beam in order to prevent from reaching the Faraday cylinder the very slow electrons and positive particles ejected from the horizontal plate on either side of the slit. The effect of the positive particles was however still detectable. The dimensions of the essential parts of the deflection chamber were as follows:

$S_1T = 10$ mm., $S_1S_2 = 80$ mm.

Source, 1 mm. × 5 mm.

Slit S_1, 4 mm. × 4 mm.

Slit S_2, 2 mm × 10 mm.

The magnetic field, of which the intensity could be calculated, was produced by means of a pair of Helmholtz coils, in such a direction as to deflect the secondary electrons as shown in Fig. 1. Electrons of any given velocity v can be focussed into the slit S_2 by adjusting the intensity of the field to the value given by

$$Hev = mv^2/\rho.$$

Actually, owing to the finite dimensions of the slits, there is a fairly sharply defined range of values of ρ, and hence of v, which enters into the current measured for a fixed value of H. This band of velocities is bounded by narrow penumbral regions *. The magnitude of this range is obviously proportional to the field intensity. Hence to obtain the true velocity curve (i.e. the number of electrons with velocities between v and $v + dv$ plotted against v, when dv is constant over the whole range of v) the ordinates must be divided by the corresponding value of v. The coils were arranged in planes perpendicular to the horizontal component of the existent magnetic field in the laboratory, the total field being used for the calculation of $H\rho$. The vertical component was neutralised by a pair of Helmholtz coils of large diameter in which the current was adjusted so as to give a zero magnetic dip at their centre. Without this compensation electrons of less than about 8 volts energy would have missed the slit S_2 entirely.

The whole of the slit system fitted closely into a brass box which was evacuated by a double stage diffusion pump through a liquid air trap. Apart from the filament and the Faraday cylinder, the whole was connected to earth. The secondary electron current to the Faraday cylinder was measured by timing the rate of charge of a Dolezalek electrometer.

* Gurney, *Proc. Roy. Soc.* vol. CIX. p. 541 (1925).

Digitized by Google

The box was pumped out for a day before any experiment and during the experiment the pumps were kept going and liquid air was on the trap continuously. A McLeod gauge on the pump side of the trap always read zero. Probably under the best conditions the pressure in the box was about 10^{-4} mm. of mercury. The total path of the secondaries was 12 cm.

Starting with zero field, readings of the current to the Faraday cylinder were taken for small increments of the intensity of the field up to a value just greater than that corresponding to the energy of the primary electrons. No measurement was made of the intensity of the primary beam. To ensure that no great change of conditions had occurred during the taking of a distribution curve, five points distributed along the curve were quickly checked afterwards and any small systematic change was corrected for.

When the magnetic field was reversed a very small negative current was still obtained which was almost independent of the field. This must be a diffuse radiation which does not describe a circular path between S_1 and S_2. It was subtracted from all readings of the secondary current. With small filament emissions this diffuse radiation was very small relative to the direct electron currents, but it grew rapidly with increasing primary currents. For this reason the smallest emissions possible were used and the electrometer was always at its full current sensitivity.

In order to make certain that the currents were due to secondary electrons two tests were made: (1) a small retarding potential was applied to the primary beam, and (2) leaving everything else as before, the target was removed and the primary beam accelerated. In neither case was any trace of current to the Faraday cylinder obtained, showing that the effect is not of photo-electric origin but is excited in the target by the primary electrons.

Results.

A series of experiments was done with a target of polished copper and primary accelerating voltages from 100 to 1000 volts. Fig. 3 shows the experimental readings, corrected for diffuse radiation, plotted for a primary energy of 123 volts. It has two peaks, the smaller of which always corresponds to an energy equal to that of the primary beam, and contains electrons which have been reflected at the target with a very small loss of energy. The small depression in the curve immediately preceding this peak is probably due to positive particles with low energies ejected from points to the right of the slit S_2 (Fig. 1) by the electrons constituting the peak focussing there. These positive corpuscles are deflected in the opposite direction to the electrons

Digitized by Google

and so reach the Faraday cylinder and produce an apparent diminution in the electron current.

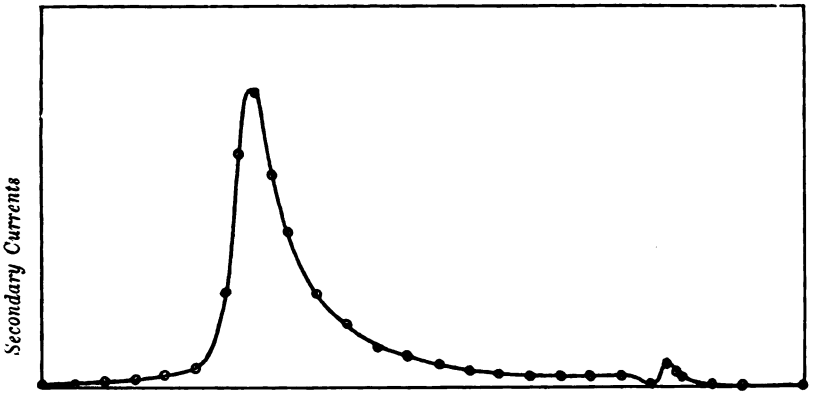


Fig. 2. Experimental curve for copper target. Primary energy 123 volts.

The position of the more intense peak is almost independent of the energy of the primary beam. As the latter increased from 123 to 1150 volts, the peak moved very slightly towards lower energies. This group obviously consists of electrons which have been ejected from the metal of the target by the comparatively fast primary electrons. Unfortunately there was no means of estimating the effect of primary energy on the intensity of the peaks, since the primary current is an unknown function of the accelerating field, and no arrangement was made for measuring it directly.

As explained above, the velocity distribution curve is deduced from the experimental curve by dividing all ordinates by the corresponding velocity. In order to obtain the energy distribution curve from this it can be shown that all abscissae must be squared and all ordinates again divided by the corresponding velocity. Fig. 3 is the result of performing these transformations on Fig. 2; it also shows the "reflected" peak magnified 100 times.

It will be noticed that the low energy side of the secondary peak indicates a very sudden rise in the number of electrons at about 10 volts, whilst above 15 volts the number falls off gradually, only becoming zero after passing the reflected electron peak.

In the range of primary velocities studied there was a slight variation of the relative intensities of the two peaks with primary velocity. For the higher velocities the reflected electrons were less prominent. This agrees with an observation of Farnsworth*

* H. E. Farnsworth, *Phys. Rev.* vol. xxv. (1925), p. 41.

Digitized by Google

in a study of the velocity distribution by the retarding potential method.

A series of experiments was done, using various substances for the target. The general form of the curves was maintained but the position of the low-energy peak varied with the target, the maximum always lying between 14 volts (for Pt) and 5 volts (for Graphite) in the order of atomic numbers as far as the experiments went.

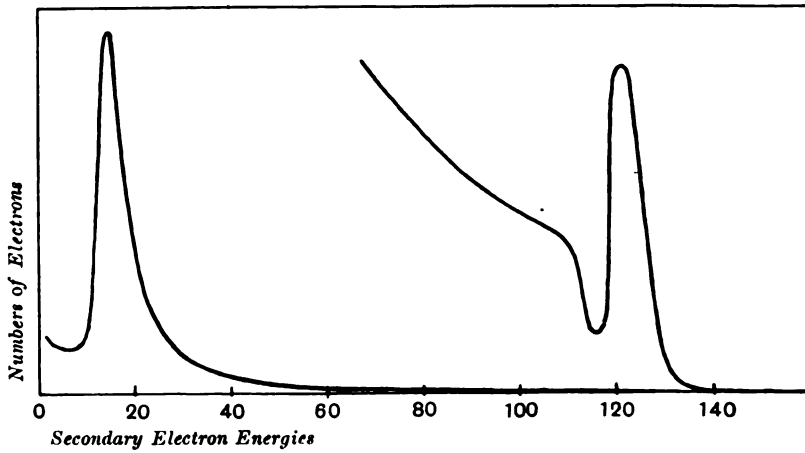


Fig. 3. Energy distribution curve.

The platinum target was arranged so that it could be heated to redness. After this treatment the target appeared to be more active but the shape of the distribution curve was unchanged, showing that the distributions obtained are not due to an adsorbed film on the surface of the target.

It was thought advisable to ascertain if the primary beam really had the energy corresponding to the applied accelerating voltage *. For this purpose the target was replaced by a filament from which the thermions were accelerated through the slit S_1. The velocity distribution was determined with the slit S_1 made of different metals. The experimental results for an aluminium slit and an energy of 420 volts are plotted in Fig. 4. Large filament emissions were used in order to show up the small secondary effects.

It is seen that most of the electrons are contained in a fairly sharp band at the energy expected. There is another group with energies as low as 5 volts, containing, in the case of the aluminium slit, relatively very few electrons. These are obviously secondaries from the edges of the first defining slit. When S_1 was of Au, Pb

* Lehmann and Osgood, *Proc. Phil. Soc. Camb.* vol. XXII. p. 731 (1925).

Digitized by Google

or Cu this group was much more prominent. As a result of this investigation the slits were made of aluminium.

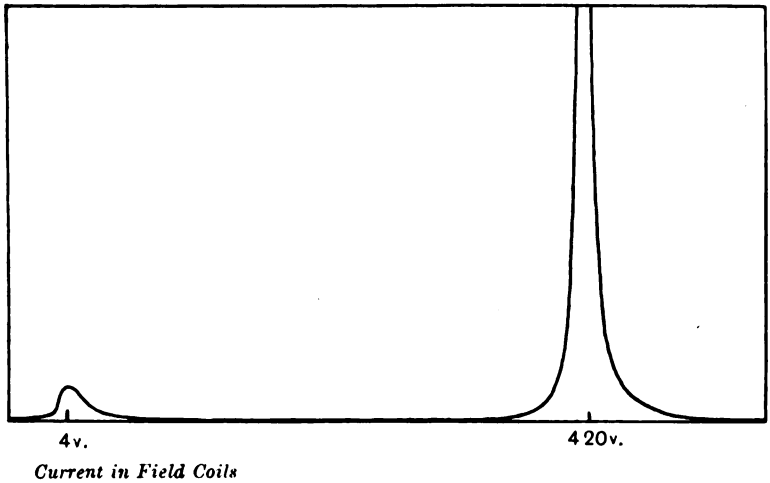


Fig. 4. Investigation of primary beam with aluminium slit. Primary energy 420 volts.

The present method confirms the previous results obtained by means of the retarding potential method in showing that the total secondary emission consists of two distinct groups of electrons with entirely different properties. The first constitutes a very small fraction of the whole (of the order of 1/100 in these experiments), is fairly homogeneous, and has an energy very nearly equal to that of the primary beam. This group must be primary electrons which have suffered close elastic collisions with atoms only, probably in the monatomic surface layer, and have not shared energy by collision with other electrons.

The second and much more intense group extends over the whole range of energies from zero up to the energy of the primary beam, showing a very prominent maximum somewhere between 5 and 14 volts according to the material of the target. Into this group would fall (*a*) any "characteristic" corpuscular radiation excited in the atoms of the target by the primary beam, and (*b*) the products of collisions between primary electrons and "free" electrons in the target. Both of these possibilities would give rise to secondary electrons with energies considerably less than the primary energy. The form of Fig. 3 might be explained on these grounds—the peak being supposed to contain the characteristic part of the secondary radiation. In his recent work, Becker infers from his retarding potential curves a Maxwellian distri-

Digitized by Google

bution with a maximum between 2 and 3 volts, for a blackened metal target. The work of Farnsworth (*loc. cit.*) also indicates a maximum at a very small energy. There is thus a considerable discrepancy between the present results and those given by the retarding potential method.

The absence of electrons with energies very nearly zero may be due to either (1) their absorption and scattering by the residual gas in the apparatus—this seems unlikely in view of the dependence of position of the peak on the material of the target; or (2) their suppression by an intense space charge just outside the surface of the target. The steepness of the low-energy side of the peak supports this latter view, but it seems doubtful whether the effect could be as large as is observed. In order to settle these points a more detailed investigation is being made under much better conditions.

In conclusion, it is a pleasure to thank Prof. Sir Ernest Rutherford for the interest he has taken in the work.

Digitized by Google

Some investigations of gas discharges by means of an exploring electrode. By Dr K. G. EMELÉUS, St John's College. (Communicated by Mr E. V. APPLETON.)

[*Received* 27 October, *read* 8 November 1926.]

1. *Introduction.*

In a glow discharge between large cold electrodes in a wide Geissler tube the main fall of potential occurs between the anode boundary of the cathode dark space and the cathode. Under the particular conditions of relatively high current density and low voltage and with the cathode dark space occupying almost the whole of the tube, the potential across the negative glow, and, if they are present, across the Faraday dark space and positive column, is small compared with the cathode fall itself. The distribution of potential in the discharge has been determined both by integration of the electric field (which can be deduced either from the deflection of a beam of cathode rays shot across the main discharge, or optically from the Stark effect), and by measuring the potential acquired by an insulated exploring electrode. Of these, the optical method has a very limited range of application, and the method based on the use of an auxiliary discharge does not permit of precise measurements outside the cathode dark space. The use of a cold exploring electrode in this way is also inexact, since the condition that the total current to it should vanish is in general not satisfied when it is at the space potential, because of the different mobilities and concentrations of the positive and negative ions in the discharge.

An alternative procedure is to find the difference of potential between the cathode boundary of the negative glow, and the anode, by Langmuir's method, in which the current to a cold exploring electrode is measured as its potential is varied*. This has been done already for glow discharges in mercury vapour†, hydrogen‡, and helium§, but in all three instances the cathode has been an incandescent filament. The present investigation is concerned with similar experiments in a glow discharge with a cold cathode. It has been carried out with a tube not built specially for the purpose, and incidental to other work, but the results which have been obtained appear to be of sufficient interest to describe separately. They illustrate the value of this method

* I. Langmuir, *Journ. Franklin Inst.*, 196, p. 751 (1923).
† McCurdy, *Phil. Mag.*, 48, p. 898 (1924).
‡ Bramley, *Phys. Rev*, 26, p. 794 (1925).
§ McCurdy and Dalton, *Phys. Rev.*, 27, p. 163 (1926).

Digitized by Google

of analysis with comparatively simple apparatus and vacuum technique, and are being followed up with another system of electrodes.

2. *Experimental Details.*

The discharge tube used is shown diagrammatically in Fig. 1. The main part of the glass wall was 4 cms. in diameter. Both anode (A) and cathode (C) were of aluminium, and the latter could be moved by the action of an external electromagnet on a cylinder of soft iron (M). The exploring electrode was a ring of platinum wire (0·362 mm. radius), sealed through the wall at opposite ends of a diameter (S): the glass round it was slightly expanded, so that the platinum effectively replaced an annular section of the main wall. The distance AS was 3·8 cms., and CS not more than 4·5 cms. The positions of the electrodes and of the edge of the negative glow were found by an external sighting device similar to that described by Aston*. The gases employed were air, and commercial oxygen, nitrogen and hydrogen, dried

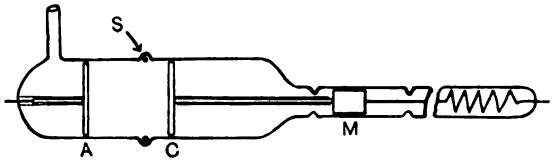


Fig. 1.

by standing over phosphorus pentoxide. Before taking measurements, the discharge tube was washed out several times with the dried gas, and then exhausted to 0·002 mm. Hg.

The electrical connections are indicated in Fig. 2, and need not be described in detail. The only points to which it is necessary to call attention are the regulation of the current from the high tension battery E by a saturated diode D, and the fact that the potential of the sound (S) was referred to that of the anode (A). F and G were shunted microammeters of low resistance, and K and L moving coil voltmeters.

The current (i) to the exploring electrode was measured as its potential (V) was changed, the difference of potential between S and the surrounding gas being localized in a sheath of ions. If S was strongly negative with respect to the surrounding space, it attracted positive ions; if it was slightly negative, it received in addition those electrons the velocity of which was sufficiently large to bring them to the electrode against the retarding field; if it was positive, electrons were attracted and positive ions

* Aston, *Roy. Soc. Proc.*, A, 79, p. 80 (1907).

Digitized by Google

repelled. The mathematical elaboration of these ideas is mainly due to Langmuir and Mott-Smith*.

With the present arrangement, the currents measured on G have to be divided by those recorded by F, to allow for slight variations in conditions of discharge during the course of a series of observations; this was usually equivalent to a maximum correction of about ten per cent. The use of this ratio has been justified by Langmuir. If the exploring electrode was made more than about 90 volts negative to the anode, the current passing between the two was liable to assume a large value discontinuously.

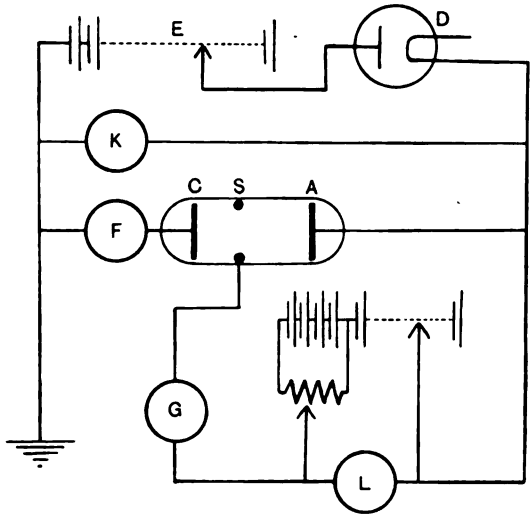


Fig. 2.

The positive ion current was sensibly linear for large collecting voltages. The electron current was obtained by extrapolating this line to small collecting voltages, and using it to correct the nett current. The logarithm of the resultant current was then plotted against the voltage, yielding in the simplest case a straight line which changed into a parabolic curve when the potential of S reached that of the surrounding space, and it began to attract electrons. The space potential was thus determined from the point of junction of the two lines. The existence of a linear relation between log i and V shows that there is a Maxwellian distribution of velocities amongst the electrons, their absolute temperature being given by the relation

$$T = 11600\, dV/d\,(\log_e i) \qquad \ldots\ldots(1),$$

* *General Electric Review*, 1924 and 1925.

Digitized by Google

or the equivalent voltage by

$$\bar{V} = T/7730 \qquad \ldots\ldots(2),$$

if V and $\bar{V}$ are in volts.

The electron concentration in the discharge may be calculated from a knowledge of the electron temperature, together with the current (I) to the collector at the space potential. On inserting the appropriate numerical values in the equation given by Langmuir and Mott-Smith, it is found that

$$n = 4{\cdot}0 \,.\, 10^{13}\, IT^{-\frac{1}{2}} \text{ per c.c.} \qquad \ldots\ldots(3),$$

if I is in amperes. It has been assumed that the collecting area was 1 sq. cm., which may be in error by 20%. This is, however, immaterial, since we are only concerned at present with relative values of n.

In certain cases, which are considered in detail in § 5, more than one group of electrons was present. No attempt has been made to investigate the positive ions. It has been found by Langmuir and others that the theory is often not applicable to collection of positive ions, and in the present experiments their presence may have been partially masked by high speed electrons which had traversed the dark space with relatively little loss of energy. The apparatus was also not suitable for visual observation of ionic sheaths.

One other point should be mentioned. When a glow discharge is not steady, the interpretation of the characteristics both of an exploring electrode and of the tube itself is uncertain*. In the present series of experiments, the discharge has been shown to be sensibly constant

(*a*) by examining it with a rotating mirror;

(*b*) by verifying the absence of a note in the earth lead;

(*c*) by the use of a heterodyne wave-meter;

(*d*) by the identity of simultaneous readings taken with intercalibrated A.C. and D.C. ammeters in series with the tube.

3. *Potential across the Cathode Dark Space.*

The measurement of this quantity was the primary object of the experiments. Some typical results are collected in Table I, in which the last column shows the difference between the potential across the tube, and the cathode fall, obtained by the procedure described in the preceding paragraph. In cases where this is negative, the anode is at a lower potential than the negative glow, and the average field between the two is the reverse of that applied to the anode and cathode. These numbers indicate the potential

* Appleton, Emeléus, and Barnett, *Proc. Camb. Phil. Soc.*, 22, p. 434 (1924); Penning, *Physica*, 6, p. 241 (1926).

Digitized by Google

immediately outside the natural positive ion sheath on the walls bounding the glow. In view of Schottky's theoretical treatment of the analogous problem of diffusion in the positive column*, it is unlikely, however, that the axis is more than one or two volts more positive.

In confirmation of earlier results, it is evident that at the lower pressures no appreciable error can result, in assuming, in calculations concerned with the cathode fall, that the whole potential across the tube is localized there. At higher pressures,

TABLE I.

Gas	Pressure mm. Hg	Tube voltage volts	Tube current milliamperes	P.D. Anode to negative glow volts
Oxygen	0·15	495	0·35	4·8
	0·19	457	0·65	5·9
Hydrogen	0·22	380	0·54	4·5
	0·22	430	0·18	−1·1
	0·39	350	1·00	34·2
Nitrogen	0·16	425	0·59	−1·1
	0·29	360	1·20	18·0
Air	0·18	528	0·50	6·5
	0·40	388	1·95	35·5

a much greater discrepancy exists; in such circumstances the only safe method is to find the space potential by Langmuir's method.

4. *Measurements in the Negative Glow and Faraday Dark Space.*

The present investigation is not strictly comparable with those mentioned in § 1, made with hot cathodes, since in the present series of experiments the exploring electrode was moved relative to the edge of the glow by an increase in the length of the discharge. The consequent changes in tube current and voltage were small, however, and were at least partially eliminated, as has been described in § 2.

The results of two sets of observations are shown in Tables II and III, in which the space potential relative to the anode, electron

* Schottky, *Phys. Zs.*, 25, p. 342 (1924).

Digitized by Google

temperature (in equivalent volts), and electron concentrations are given for various positions of the exploring electrode. They are typical of a large number which have been made in hydrogen and in the other gases.

A striking fact is that at the lower pressure, the cathode edge of the negative glow was at a higher potential than the anode.

TABLE II.

Distance from edge of cathode D.S towards anode	0·3	3·8	4·5	4·7	9·0	12·9	mms.
Space potential	34·2	36·0	36·3	36·5	35·9	36·0	volts
Electron temperature and concentration	8·8	8·2	7·6	7·0	—	—	volts
Fast group	4·85	5·67	5·65	5·74	—	—	10^7 per c.c.
Electron temperature and concentration	4·7	2·7	3·0	2·2	4·5	4·7	volts
Slow group	2·47	8·77	5·87	4·61	11·38	7·65	10^7 per c.c.
Total electron concentration	7·32	14·44	11.52	10·35	11·38	7·65	10^7 per c.c.

Gas—Hydrogen. Pressure—0·39 mm. Hg.
Tube current—1·0 mA. Tube voltage—350 volts.
Thickness of cathode dark space—3·5 mms.

This exaggerated reversal of the field is of frequent occurrence in low voltage arcs, and is found in a form of glow discharge in which anode and cathode are at the same potential*, but was not present in the hot cathode glows in either hydrogen, mercury or helium. It cannot be decided yet whether it is characteristic of the cold cathode, or of the small current density, of the order of 0·1 mA./sq. cm.† It is also seen that at the lower pressure, the space

* Hufford, *Phil. Mag.*, 50, p. 1197 (1925).

† Usually the negative glow was the only bright part of the discharge.

Digitized by Google

potential diverged more from that of the anode as the latter was approached, the apparent reversed field being greatest at the negative edge of the glow. At the higher pressure, the potential was sensibly constant at about 36 volts, except in the glow, where it was again definitely a little lower. This effect of pressure is in the same sense as was observed by Bramley in hydrogen, the field in the Faraday dark space being reversed only at low pressures.

TABLE III.

Distance from edge of cathode D.S. towards anode	0·2	3·9	4·4	11·5	20·9	mms.
Space potential	−1·1	4·6	8·2	11·0	32·3	volts
Electron temperature and concentration	17·3	9·0	—	—	—	volts
Fast group	1·69	4·97	—	—	—	10^7 per c.c.
Electron temperature and concentration	3·5	3·9	5·6	6·8	5·7	volts
Slow group	5·26	6·18	9·17	8·07	3·77	10^7 per c.c.
Total electron concentration	6·95	11·15	9·17	8·07	3·77	10^7 per c.c.

Gas—Hydrogen. Pressure—0·2 mm. Hg.
Tube current—1·4 mA. Tube voltage—430 volts.
Thickness of cathode dark space—8·2 mms.

Additional confirmation of its reality in the present experiments is furnished by the marked diminution either in the total electron concentration, or in the temperature of the electrons as the anode is approached; this allows a current to flow by diffusion against the electric field.

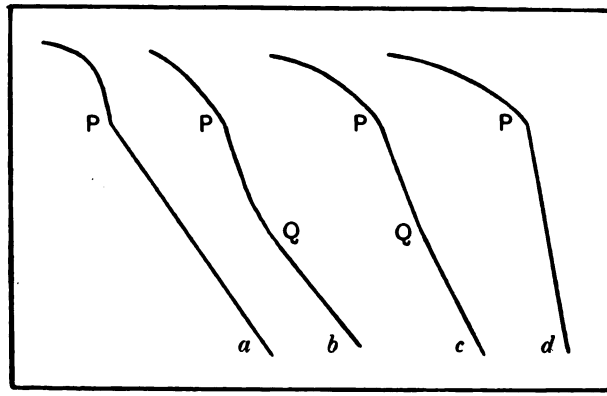
In several instances it was found that, after an initial decrease, a slight increase in electron concentration again took place on approaching the anode: this was probably due to the existence of a positive striation, too feebly luminous to be detected visually.

Digitized by Google

5. *Electron Temperatures.*

The detail of the curves obtained when the logarithm of the corrected electron current is plotted against the collecting voltage on the exploring electrode promises to yield a great deal of information. Four main types of such curves can be distinguished, which appear to be characteristic of different parts of the discharge.

If the sound is situated in that part of the negative glow which is immediately adjacent to the cathode dark space, and which is often less bright than the remainder, the curve takes the form shown in Fig. 3(*a*). The space potential is indicated by the break *P*, and a large increase in the current takes place immediately electrons are attracted. The linear part has a small slope, indicating a Maxwellian distribution of velocities amongst



Ordinates: logarithm of electron current.
Abscissae: potential of exploring electrode.

Fig. 3.

electrons of high temperature, often of the order of 100,000° K. (13 volts). These two facts are probably connected. Swift electrons give rise to slow secondary electrons when incident on the exploring electrode, and the latter are rejected below the space potential, the semi-log curve being moved laterally without change of slope. On the other hand, most of these electrons are retained immediately above the space potential. This simple type of curve was not obtained in either of the runs, the data for which are given in Tables II and III, but this is believed to be due to the region of the discharge of which it is characteristic being very narrow. When it was present, the electron concentration was always much less than it was a few mms. farther into the glow, in spite of the presence of a reversed field. This may be attributed

Digitized by Google

partly to the compensating effect of the large increase in temperature, but partly also to the secondary electrons reducing the recorded current at the space potential, and hence the calculated concentration (eq. 3). The relative importance of these two factors might be decided by using two collectors, identical except in the nature of the metal employed.

The second stage is shown in Fig. 3(*b*), and is characteristic of the cathode side of the brightest part of the negative glow. For large retarding voltages there is, as in (*a*), a line of small slope. At some point Q this starts to bend upwards, again changing in slope more or less abruptly at the space potential P. The existence of the region PQ is interpreted to show the presence of a second group of slower electrons. Their temperature and concentration are found by allowing for the fast group by extrapolating the linear part of the curve to smaller retarding voltages, and then plotting a fresh semi-log curve for the difference of currents. This has been done in obtaining the data for the two groups of electrons in Tables II and III. It is not possible to obtain accurate information about the slow electrons, because the difference of currents is usually quite small except for a small region just below the space potential. When these two groups are present, the localization of the break P is also a little uncertain, but values obtained from the curve (*b*) usually agreed with those found from the semi-log curve for the slow electrons to within 1·5 volts. On reference to Table II, it will be seen that the average energy of the fast group falls off as the anode is approached, but that their concentration is almost constant, except on the extreme cathode side of the glow, where secondary effects similar to those discussed in the last paragraph probably come into play. There is insufficient evidence to show how the slower group changes.

The third stage (Fig. 3*c*) is found in the anode side of the negative glow, and in neighbouring parts of the Faraday dark space. Two groups of electrons are again present, but the difference which would be obtained by the method adopted in analyzing (*b*) is now so small that it is best to treat them as if they were homogeneous, with a temperature determined by the average slope.

The last stage (*d*) is that originally contemplated by Langmuir, when the electrons have one definite and usually fairly low temperature. In the present experiments this is characteristic of the Faraday dark space.

Bramley has recorded curves similar to (*b*) in hydrogen at low pressures, close to the hot cathode, and Langmuir and Mott-Smith mention the occurrence of curves of type (*c*) in arcs in mercury. Langmuir has also studied low current arcs in mercury vapour at pressures of the order of 0·25 bar, using a

Digitized by Google

hot cathode, and has distinguished three classes of conducting electrons*.

The "primary electrons" are a group with a small random component of velocity superposed on that acquired by a free fall through a potential difference slightly smaller than that applied to the tube. It is unlikely that these occur in a cold cathode discharge, because the mean free path of an electron is considerably less than the thickness of the cathode dark space. The primary electrons might be replaced by a group having a random distribution about a terminal velocity acquired in the cathode dark space. Indications of the existence of these have been found for retarding potentials of the order of 75 volts, but the discharge then tends to be unstable (§ 2).

The "ultimate electrons" are of low average velocity, and are probably in thermal equilibrium with radiation and the positive ions. In the present experiments, these occur in (*d*), and as the slow component in (*b*) and (*c*).

The "secondary electrons" are intermediate between the other types. They may be identified with the fast components in (*b*) and (*c*), and in view of the continuous transformation from (*b*) to (*a*), it appears that on the cathode edge of the negative glow, they are present unaccompanied by the other groups. Insufficient data are available to enable us to decide if the secondary and ultimate electrons in the negative glow approximate in temperature to form the ultimate electrons in the Faraday dark space.

A number of observations have been made with the exploring electrode in the cathode dark space. The currents to the exploring electrode are small, and curves of a totally different nature are obtained. In this region, Langmuir's theory is not directly applicable, because of the presence of fast positive ions and soft X-rays, and it will not be considered further at present.

As has already been pointed out, the results described in the two preceding sections are to a certain extent ambiguous, because the cathode has been moved, and a large exploring electrode used. There is also a possibility that curves of types (*a*), (*b*) and (*c*) are affected by radiation. Further experiments are in progress, with a discharge tube of different design.

6. *Summary.*

The glow discharge between cold aluminium electrodes in air, oxygen, nitrogen and hydrogen has been analyzed by Langmuir's method, for pressures between 0·1 and 0·4 mm. Hg, current densities of from 0·02 to 0·2 mA./sq. cm., and applied potentials between 300 and 700 volts. An annular exploring electrode has

* I. Langmuir, *Phys. Rev.*, 26, p. 585 (1925).

Digitized by Google

been used. It has been found that whilst practically the whole fall of potential is localized across the cathode dark space at the lower pressures, a fall of as much as 40 volts can exist across the remainder of the discharge at the higher pressures. Reversal of the electric field has been found in the negative glow, and in certain cases in the Faraday dark space, when conditions are favourable for passage of an electron current by diffusion against the field. In several instances the negative glow was at a higher potential than the anode. Two groups of electrons occur in the negative glow, together with a single fast group at the anode boundary of the cathode dark space, and a single slow group in the Faraday dark space.

The experiments described have been carried out in the Wheatstone Laboratory at King's College, University of London. The author is indebted to Professor E. V. Appleton for his interest and advice during the course of the research, as well as for his assistance in making some of the observations.

Digitized by Google

The calculation of atomic fields. By L. H. THOMAS, B.A., Trinity College.

[*Received* 6 November, *read* 22 November 1926.]

The theoretical calculation of observable atomic constants is often only possible if the effective electric field inside the atom is known. Some fields have been calculated to fit observed data* but for many elements no such fields are available. In the following paper a method is given by which approximate fields can easily be determined for heavy atoms from theoretical considerations alone.

1. *Assumptions and the deduction from them of an equation.*

The following assumptions are made.

(1) Relativity corrections can be neglected.

(2) In the atom there is an effective field given by potential V, depending only on the distance r from the nucleus, such that

$$V \to 0 \text{ as } r \to \infty,$$

$$Vr \to E, \text{ the nuclear charge, as } r \to 0.$$

(3) Electrons are distributed uniformly in the six-dimensional phase space for the motion of an electron at the rate of two for each h^3 of (six) volume. (This means one for each unit cell in the phase space of translation and rotation of a spinning electron.) The part of the phase space containing electrons is limited to that for which the orbits are closed.

(4) The potential V is itself determined by the nuclear charge and this distribution of electrons.

In reality the effective field at any point depends on whether the point is empty or occupied by a foreign electron or one or another atomic electron and on the circumstances of that occupation. These fields can only be expected to be sensibly the same or approximately calculable from the above assumptions if the density of electrons is large, that is, in the interior of heavy atoms.

If e, m, p are the charge, mass and momentum of an electron, the Hamiltonian function for the electronic motion is ((1) and (2) above),

$$\frac{1}{2m} p^2 - eV.$$

* D. R. Hartree, *Proc. Camb. Phil. Soc.*, **21**, p. 625; E. Fues, *Zeit. für. Phys.*, **11**, p. 369.

Digitized by Google

There are ((3) above) electrons at two for each h^3 of phase space for which

$$p < (2meV)^{\frac{1}{2}}$$

i.e. at

$$\frac{2}{h^3}\frac{4}{3}\pi(2meV)^{\frac{3}{2}}$$

per unit of ordinary (coordinate) space.

Thus ((4) above)

$$\nabla^2 V = 4\pi e \cdot \frac{2}{h^3}\frac{4}{3}\pi(2meV)^{\frac{3}{2}}$$

i.e.
$$\frac{1}{r^2}\frac{d}{dr}r^2\frac{dV}{dr} = 4\pi e\frac{2}{h^3}\frac{4}{3}\pi(2me)^{\frac{3}{2}}V^{\frac{1}{2}} \qquad \text{.....(1·1),}$$

with ((2) above)

$$V \to 0 \text{ as } r \to \infty,$$
$$Vr \to E \text{ as } r \to 0.$$

Now express distance in terms of the 'radius of the normal orbit of the hydrogen atom,' $a = h^2/4\pi^2 me^2 = 5{\cdot}3 \cdot 10^{-9}$ cms., potential in terms of the potential of an electron at this distance, so

$$r = \rho\frac{h^2}{4\pi^2 me^2},$$

$$V = \psi e \Big/ \frac{h^2}{4\pi^2 me^2},$$

and equation 1·1 becomes

$$\frac{1}{\rho^2}\frac{d}{d\rho}\left(\rho^2\frac{d\psi}{d\rho}\right) = \frac{8\sqrt{2}}{3\pi}\psi^{\frac{3}{2}} \qquad \text{.....(1·2),}$$

with
$$\psi \to 0 \text{ as } \rho \to \infty,$$
$$\rho\psi \to N, \text{ the atomic number, as } \rho \to 0.$$

(It is useful to note that with 'a' as unit of length, the charge and mass of the electron as units of charge and mass, $h = 2\pi$, whence 1·2 is at once verified.)

The 'effective nuclear charge' at distance ρ is then given by

$$Z = -\rho^2\frac{d\psi}{d\rho}.$$

Putting $\psi = \frac{9\pi^2}{128}\phi$, the equation for ϕ is

$$\frac{1}{\rho^2}\frac{d}{d\rho}\left(\rho^2\frac{d\phi}{d\rho}\right) = \phi^{\frac{3}{2}} \qquad \text{.....(1·3).}$$

Digitized by Google

2. *Discussion of the equation.*

Write $\log \rho = x$, $\rho^{\frac{1}{2}}\phi = w$, and the equation becomes

$$\frac{d^2w}{dx^2} - 7\frac{dw}{dx} + 12w = w^{\frac{3}{2}} \qquad \text{.....(2·1),}$$

or, if $dw/dx = p$

$$\frac{dp}{dw} = 7 + \frac{w(w^{\frac{1}{2}} - 12)}{p} \qquad \text{.....(2·2).}$$

The maximum and minimum locus of this equation is

$$p = -\frac{w(w^{\frac{1}{2}} - 12)}{7}.$$

The inflexion locus is

$$p = -\frac{2w(w^{\frac{1}{2}} - 12)}{7 \mp (1 + 6w^{\frac{1}{2}})^{\frac{1}{2}}} = f(w),$$

and

$$\left(\frac{dp}{dw}\right)_{p=f(w)} - f'(w) = \pm 3w^{\frac{1}{2}}(w^{\frac{1}{2}} - 12)\{7 \mp (1 + 6w^{\frac{1}{2}})^{\frac{1}{2}}\}^{-2}(1 + 6w^{\frac{1}{2}})^{-\frac{1}{2}},$$

gives the direction in which the solutions cross the inflexion locus.

There are two singular points, $w = 0$, $p = 0$; $w = 144$, $p = 0$.

At $w, p \to 0$, $(4w - p)^4 \sim c(p - 3w)^3$, $(w > 0)$,

at $w \to 144$, $p \to 0$, $(7{\cdot}772(w - 144) - p)^{7{\cdot}772}(7{\cdot}772(w - 144) + p) \sim c$,

give the form of the solutions, c being arbitrary.

The dp/dw discriminant gives $p = 0$, and $w = 144$ or $\phi = 144\rho^{-4}$ as a singular solution.

There is an approximate particular solution,

$$\left.\begin{aligned} p &= -\frac{4\lambda}{\sqrt{12}} w(w^{\frac{1}{2}} - 12^{\frac{1}{2}}) \\ w &= \frac{144}{(1 + e^{-\lambda x})^4} \end{aligned}\right\} \qquad \text{.....(2·21),}$$

which satisfies

$$\frac{3}{5\lambda}\frac{dp}{dw} = \frac{12}{5} + \frac{3}{\lambda} + \frac{w(w^{\frac{1}{2}} - 12)}{p}.$$

The solutions of 2·2 lie roughly as in the sketch (Fig. 1), the arrows give the direction of increase of ρ. The only solutions for which $\phi \to 0$ as $\rho \to \infty$ and $\phi = O(1/\rho)$ as $\rho \to 0$ correspond to the solution through O and A in the sketch—2·21 is an approximation to this solution*. Different values of the nuclear charge

* It is only at A that ρ becomes infinite.

Digitized by Google

correspond to the replacement of x by $x+c$, which does not affect 2·1, so that if the equation is integrated numerically, starting from an initial position with w and p near A and any value of x, all the required solutions can be deduced.

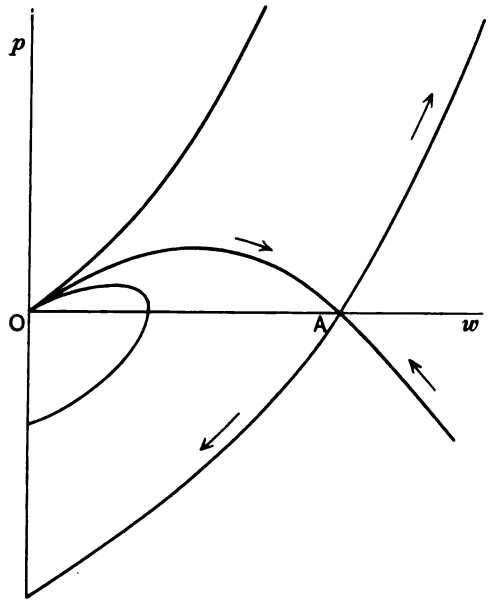


Fig. 1

3. *The numerical integration.*

For the initial values put

$$w = 144\,(u+1)^{-4}, \quad v = \frac{du}{dx},$$

in 2·1, obtaining

$$(u+1)\,v\,\frac{dv}{du} - 5v^2 - 7\,(u+1)\,v - 3u\,(u+2) = 0.$$

If $v' = \lambda\,(u + au^2)$,

$$G \equiv (u+1)\,v'\,\frac{dv'}{du} - 5v'^2 - 7\,(u+1)\,v' - 3u\,(u+2)$$

$$= u\,(\lambda^2 - 7\lambda - 6) + u^2\,(-4\lambda^2 - 7\lambda - 3 + a\,(3\lambda^2 - 7\lambda))$$
$$+ u^3\,(a\,\{-7\lambda^2 - 7\lambda\} + 2a^2\lambda^2) - 3a^2\lambda^2u^4$$

$$= \tfrac{1}{2}u^2\,\{35\sqrt{73} - 292 + a\,(134 - 14\sqrt{73})\}$$
$$+ au^3\,\{7\,(4\sqrt{73} - 34) + a\,(61 - 7\sqrt{73})\} - 3a^2\lambda^2u^4.$$

Digitized by Google

For $\lambda = -\frac{1}{2}(\sqrt{73} - 7) = -\cdot 77200$ and $u > 0$.

$G < 0$ for $a = -(35\sqrt{73} - 292)/(134 - 14\sqrt{73}) = -\cdot 0027900$,

$G > 0$ for $a = 0$,

from which it can be shown that for $u > 0$

$$-\cdot 77200u < \frac{du}{dx} < -\cdot 77200(u - \cdot 0027900u^2) \quad \ldots(3\cdot 1).$$

For the actual numerical integration it is convenient to put

$$x = X \log_e 10, \quad w = 144 . 10^Y,$$

so that 2·1 becomes

$$\frac{d^2Y}{dX^2} = \log_e 10 \left\{12 . 10^{\frac{1}{2}Y} + \cdot 25 - \left(3\cdot 5 - \frac{dY}{dX}\right)^2\right\} \quad \ldots(3\cdot 2),$$

while

$$\psi = \frac{9\pi^2}{128} 144 . 10^{Y - 4(X+c)} \quad \ldots\ldots(3\cdot 31),$$

$$Z = \frac{9\pi^2}{128}\left(4 - \frac{dY}{dX}\right) 144 . 10^{Y - 3(X+c)} \quad \ldots\ldots(3\cdot 32),$$

$$\rho = 10^{X+c} \quad \ldots\ldots(3\cdot 33),$$

where c is to be determined from the atomic number. Z is the effective nuclear charge.

If $Y = -1, \quad u = 10^{\frac{1}{2}} - 1 = \cdot 77828,$

and $$1\cdot 3515 > \frac{dY}{dX} > 1\cdot 3489 \qquad \text{(from 3·1)}.$$

Starting with

$$X = 0, \quad Y = -1, \quad \frac{dY}{dX} = 1\cdot 35,$$

numerical integration was carried out by steps of ·1 to $X = -3$ by the aid of the formulae

$$\left(\frac{dy}{dx}\right)_{n+1} - \left(\frac{dy}{dx}\right)_n = \left(\frac{d^2y}{dx^2}\right)_n + \frac{1}{2}\Delta\left(\frac{d^2y}{dx^2}\right)_{n-1} + \frac{5}{12}\Delta^2\left(\frac{d^2y}{dx^2}\right)_{n-2} + \ldots,$$

$$y_{n+1} - y_n = \left(\frac{dy}{dx}\right)_n + \frac{1}{2}\left(\frac{d^2y}{dx^2}\right)_n + \frac{1}{6}\Delta\left(\frac{d^2y}{dx^2}\right)_{n-1} + \frac{1}{8}\Delta^2\left(\frac{d^2y}{dx^2}\right)_{n-2} + \ldots *.$$

* See Whittaker and Robinson, *The Calculus of Observations*, p. 365.

Digitized by Google

For $X = -3$ it appears that

$$3{\cdot}5 - \frac{dY}{dX} = {\cdot}508, \quad \log_{10} 144 + Y = \bar{7}{\cdot}4385$$

so equation 3·32 gives

$$Z = \frac{9\pi^2}{128} \cdot 1{\cdot}008 \cdot 10^{2{\cdot}4385 - 3c},$$

i.e.
$$c = {\cdot}7611 - \tfrac{1}{3} \log_{10} N,$$

since here closely enough $Z = N$ the atomic number.

e.g. for $N = 55$ (caesium), $c = {\cdot}1810$.

4. *Numerical results.*

The following table gives the values of

$$3{\cdot}5 - \frac{dY}{dX} \text{ and } \log_{10} 144 + Y$$

found by numerical integration and the corresponding values of ρ, Z, ψ for caesium. The former may be in error by about 10 in the last decimal place.

For $\rho_0 < {\cdot}006$, the field is sensibly a Coulomb field.

For $\rho_0 > 1{\cdot}5$, the approximate formula 2·21 is an accurate enough solution of the differential equation, but this equation is not an accurate representation of the facts.

For the element of atomic number N the corresponding values are given by

$$\rho = \rho_0 \left(\frac{55}{N}\right)^{\frac{1}{3}},$$

$$Z = Z_0 \left(\frac{N}{55}\right),$$

$$\psi = \psi_0 \left(\frac{N}{55}\right)^{\frac{4}{3}}.$$

The values Z_1 are (unpublished) values calculated by Mr Hartree for caesium from the observed levels and which he has very kindly allowed me to include for comparison.

In conclusion, I wish to thank Professor Bohr and Professor Kramers for their encouragement when I was carrying out the numerical integration last March.

Digitized by Google

$-X$	$3{\cdot}5-\frac{dY}{dX}$	$\log_{10}144+Y$	ρ_0	Z_0	ψ_0	Z_1
0	2·150	1·1584	1·517	7·6	1·887	9·9
·1	2·015	1·0167	1·205	10·4	3·412	12·5
·2	1·880	·8614	·9572	13·7	6·008	16·0
·3	1·746	·6927	·7603	17·5	10·23	19·7
·4	1·615	·5105	·6040	21·6	16·90	24·3
·5	1·489	·3156	·4800	25·8	27·10	29·0
·6	1·371	·1086	·3811	30·1	42·26	33·4
·7	1·261	$\bar{1}$·8901	·3027	34·2	64·18	36·6
·8	1·160	$\bar{1}$·6611	·2404	38·0	95·15	39·5
·9	1·069	$\bar{1}$·4225	·1910	41·3	138·0	42·2
1·0	·987	$\bar{1}$·1752	·1517	44·2	196·1	44·7
1·1	·914	$\bar{2}$·9202	·1205	47·7	273·9	46·6
1·2	·851	$\bar{2}$·6584	·09572	48·7	376·4	47·8
1·3	·795	$\bar{2}$·3906	·07603	50·3	510·4	48·4
1·4	·747	$\bar{2}$·1176	·06040	51·5	683·8	49·3
1·5	·706	$\bar{3}$·8402	·04800	52·5	906·8	50·6
1·6	·671	$\bar{3}$·5590	·03811	53·2	1198	51·6
1·7	·642	$\bar{3}$·2746	·03027	53·8	1556	52·4
1·8	·614	$\bar{4}$·9875	·02404	54·0	2018	53·4
1·9	·595	$\bar{4}$·6979	·01910	54·4	2601	53·9
2·0	·577	$\bar{4}$·4064	·01517	54·6	3340	54·1
2·1	·564	$\bar{4}$·1134	·01205	54·8	4273	54·4
2·2	·552	$\bar{5}$·8191	·009572	54·9	5450	54·6
2·3	·542	$\bar{5}$·5238	·007603	54·9	6936	54·7
2·4	·534	$\bar{5}$·2276	·006040	55·0	8809	54·8
2·5	·527	$\bar{6}$·9306	·004800	55·0	11170	
2·6	·521	$\bar{6}$·6330	·003811	55·0	14140	
2·7	·517	$\bar{6}$·3349	·003027	55·0	17870	
2·8	·513	$\bar{6}$·0364	·002404	55·0	22580	
2·9	·510	$\bar{7}$·7376	·001910	55·0	28570	
3·0	·508	$\bar{7}$·4385	·001517	55·0	35960	

Digitized by Google

The Tensile Deformation of Large Aluminium Crystals at Crystal Boundaries. By R. L. ASTON, Barker Graduate Scholar of the University of Sydney, Dominion Exhibitioner of Trinity College. (Communicated by Mr G. I. TAYLOR.)

[*Received* 11 November 1926, *read* 19 January 1925.]

(Plate IV.)

The tensile deformation of single crystals of aluminium was examined by Taylor and Elam, using test-pieces composed of one crystal only. Miss Elam kindly supplied the author with a test-piece containing three large crystals, and an investigation of the effect of the constraint at the crystal boundary was carried out upon it as described below.

Taylor and Elam found* that plastic deformation takes place by a general shear on planes parallel to a (111) type of crystal plane and in the direction of a [110] type of axis lying in that plane. The particular plane and axis chosen depend upon the orientation of the crystal relative to the axis of stress and are those for which there is the greatest component of stress producing shear. If a stereographic projection of the crystal axes be divided into triangles, each with a [100], a [110] and a [111] type of axis as its vertices, the point representing the axis of stress will be seen to move, after deformation of the crystal, along a line crossing the [100]–[111] side of the triangle in which it lies, toward the diad axis next in that direction. After the point reaches the side of the triangle, slipping takes place upon two sets of planes and the point moves along the side toward the neighbouring [211] type of axis until fracture occurs.

Tests of the deformation of aluminium specimens containing more than one large crystal were made by Carpenter and Elam on flat bars, such as were used in their earlier experiments on crystal growth, but there were no X-ray determinations of the axes. They found† that the individual crystals deformed in the normal way, but there were marked modifications at the boundaries. There appeared to be no slipping along a boundary and fracture never occurred there: but "the reduction both in thickness and in breadth was much less at this point" and "the influence of one crystal in giving support to the other...was often noticeable at a distance of 0·25 inch from the boundary."

The specimen used by the author was a round aluminium test-piece machined to a diameter of 0·564 inch from a one inch diameter bar. The parallel portion was made up of three large

* *Roy. Soc. Proc.* A, **102** (1923), p. 643; **108** (1925), p. 28.

† *Roy. Soc. Proc.* A, **100** (1921), p. 329; *Journ. Iron and Steel Inst.*, No. I, 1923, p. 175.

Digitized by Google

crystals, grown in it by Carpenter and Elam, of which one occupied one end of the bar while the other end was occupied by two longitudinally-adjacent crystals. The traces of the crystal boundaries are shown in the development drawn in Fig. 1, which also shows the eight reference lines, which were ruled as light scratches on the surface. The crystals were distinguished by the letters "A." "B," and "C" respectively.

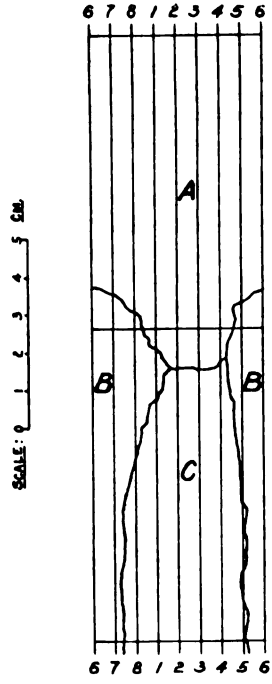


Fig. 1. Before extension.

An X-ray analysis of the orientation of each crystal was made before the test-piece was extended. It was then subjected to an elongation of about 10 % in the tensile-testing machine used and described by Farren and Taylor*. A series of analyses was then made in the vicinity of the crystal boundaries, along reference line (1). More analyses were made after a further extension to about 20 % elongation.

The analysis was carried out by the method devised by Müller† for the work of Taylor and Elam. But it was now complicated by the fact that the specimen, not being a single-crystal,

* *Roy. Soc. Proc.* A, **107** (1925), p. 427.
† *Roy. Soc. Proc.* A, **105** (1924), p. 500.

Digitized by Google

no longer extended uniformly. Thus the longitudinal lines ruled on the surface to establish the reference planes did not remain straight throughout the extension but were subject to large deviations and to intense local distortions. The particles which originally lay in a plane of reference might come to lie in quite other directions than that of the surface, joining corresponding points of the two reference lines, by which the plane was originally defined. Thus, the origin of azimuthal co-ordinates is no longer

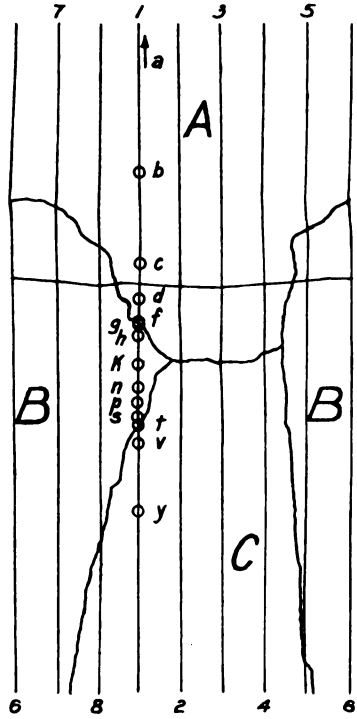


Fig. 2. Extension "10 %."

accurately known but can only be assumed at any point still to be directed toward the corresponding point of the companion line. The error which is consequently involved in the value of the azimuthal co-ordinate will have no effect upon a comparison of axes of the same crystal measured at the same place at the same extension, but will introduce errors into comparisons of the azimuthal co-ordinates for axes measured at different places or at different extensions, since the unknown correction will, in general, have altered. Thus, in Crystal "B" at "10 %" elongation there

Digitized by Google

was a spurious change in the azimuthal co-ordinates of axes measured at points along line (1) which amounted to over 10° in a distance of not more than 12 mm.

It is thus impossible to detect any azimuthal change of orientation in a crystal. It was decided therefore to follow the movement of the axis of the specimen through the crystal axes, as described above. The effect of unknown and varying errors in azimuthal co-ordinates is then no longer relevant, since it concerns a rotation of the crystallographic diagram about an axis and this does not alter the position of the axis in the diagram.

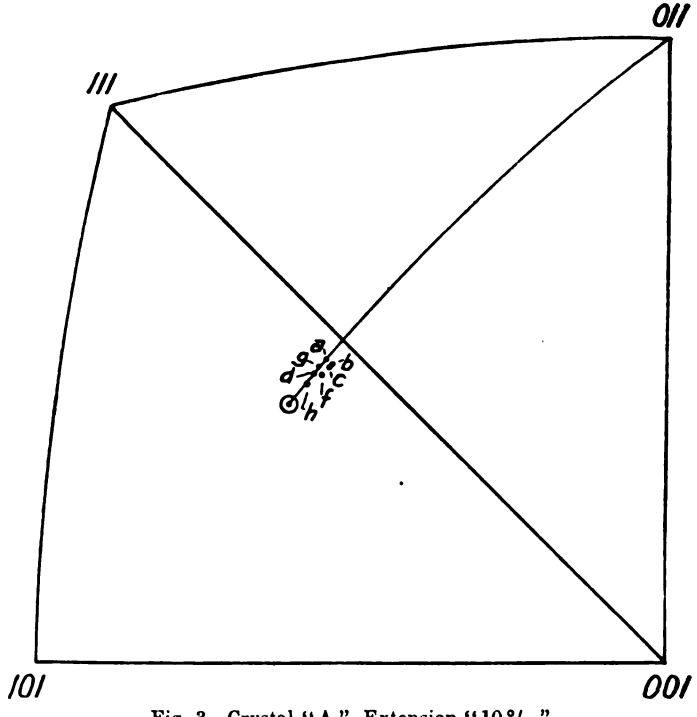


Fig. 3. Crystal "A." Extension "10%."

The appearance of the surface of the test-piece after deformation showed effects similar to those observed by Carpenter and Elam*. Crystal "A," away from the region of the boundary, deformed in the normal way, the cross-section reducing to an ellipse with the major axis approximately in the plane (4·3)–(8·3)†. Crystal "B" had an orientation rather similar to that of Crystal

* *Loc. cit.*

† This notation indicates that the extremities of the major axis of the elliptical cross-section lay between reference lines 4 and 5 and between lines 8 and 1.

Digitized by Google

"C," and the two together deformed into an approximately elliptical cross-section, with major axis in the plane (3·8)–(7·8). In the centre of the bar, where Crystals "B" and "C" abutted against Crystal "A," the reduction of the cross-section in area and the departure from a circular shape were less than in any other portion. Nevertheless the general trend of deformation peculiar to each crystal persisted in this region and the differences in surface contour were adjusted by steep changes of slope in the immediate neighbourhood of the boundaries. Observing the crystal boundaries

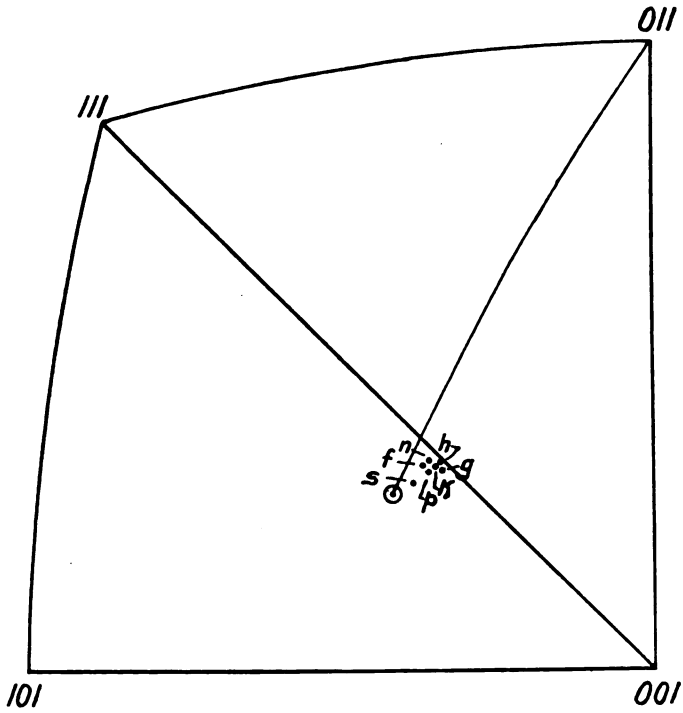


Fig. 4. Crystal "B." Extension "10 %."

at points where a reference line crossed them, no slip of one crystal over another could be detected with the aid of a low-power microscope. Even for extreme deformations, to which this and other similar specimens were afterwards submitted, the lines were unbroken in passing from one crystal into an adjacent one. The direction of the reference lines was, however, considerably modified near to the crystal boundaries, as is shown in Figs. 2 and 6, which have been drawn for the "10 %" and "20 %" elongations respectively. It was evident that the deformation at the boundary was

Digitized by Google

small; the diminution of strain extended into each crystal for two or three millimetres, while some of the longitudinal reference lines showed a deviation from their normal direction for a distance of as much as five centimetres from crossing a transverse boundary.

These observations of the rigidity of the boundary and the gradual accommodation in the crystal were supported by the X-ray analyses. The points at which analyses were made have been marked by circles in Figs. 2 and 6 and each has been given a distinguishing letter. The results of the analyses have been shown by

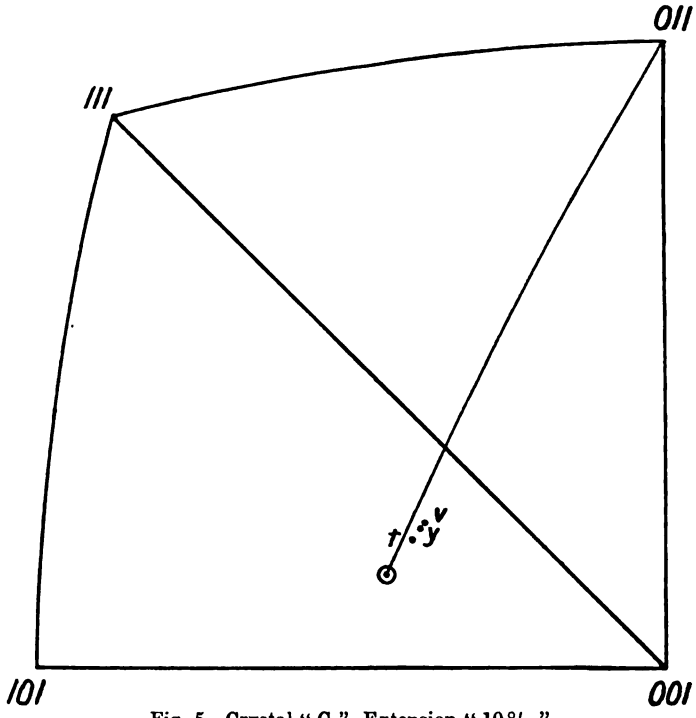


Fig. 5. Crystal "C." Extension "10 %."

plotting the position of the axis of the specimen relative to the crystal axes, the stereographic diagrams shown in Figs. 3 to 5 for "10 %" elongation and in Figs. 7 to 9 for "20 %" elongation. These diagrams show a portion of the complete stereographic representation of the axes of a cubic crystal. The axis (100) is in the centre of the projection and the points to which the letters *a*, *b*, *c*, etc.* are attached represent the direction of the axis of the specimen relative to the crystal axes. The corresponding distin-

* See Taylor and Elam, *Roy. Soc. Proc.* **102** (1923), p. 643 and **108** (1925), p. 28.

Digitized by Google

guishing letter has been placed alongside of each point. The relative position of the axis before extension is also shown in each diagram, surrounded by a ring. The variations in orientation at different points are not large compared with the order of accuracy of the analysis, which is subject to errors of as much as one degree, but the nature of the variations is nevertheless definite and consistent. While in each case the axis has tended to move toward the [011] crystal axis indicated by the work of Taylor and Elam,

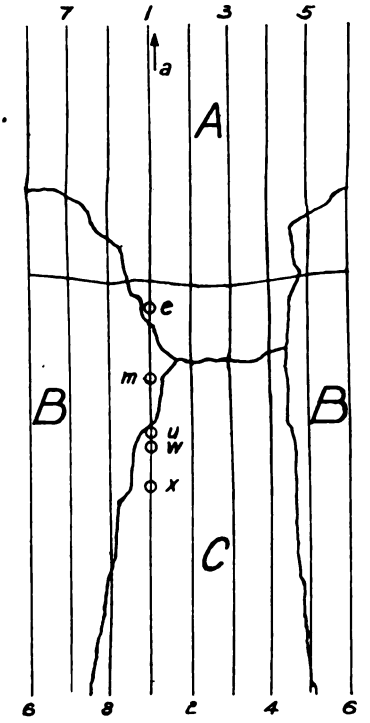


Fig. 6. Extension "20 %."

the extent to which the axis has moved varies with the distance from the boundary.

Thus, in Crystal "A" at "10 %" elongation, Fig. 3, the points marked *a*, *b* and *c*, which are all from positions well within the body of the crystal, show about the same change of orientation. But *d*, from a position within 2 mm. of the boundary, and *f* and *g*, still closer to the boundary, show a smaller movement of the axis. For the point *h*, reflections were obtained corresponding to Crystal "A" when the centre of the incident beam was about a millimetre

Digitized by Google

across the boundary into Crystal "B." These would be from the small slope on the edge of Crystal "A" immediately to the right, which would be hit by a portion of the primary beam. They indicate a deformation which is no more than a half of that found in the body of the crystal.

Similarly, in Crystal "B," Fig. 4, the points *h*, *k* and *n*, representing analyses in positions between the two boundaries, show a greater change of orientation than the point *f*, at one boundary, and a decidedly greater change than the point *s*, at the other

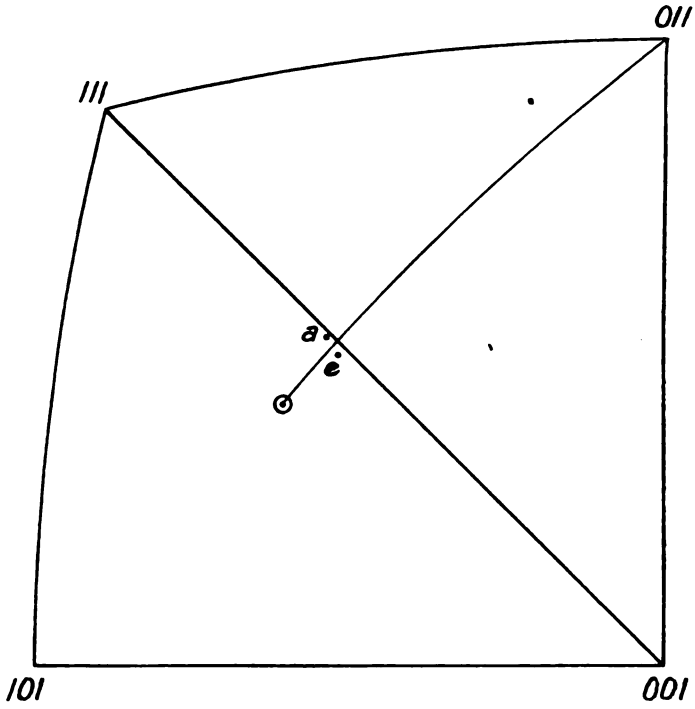


Fig. 7. Crystal "A." Extension "20 %."

boundary. Again, in Crystal "C," Fig. 5, the points *v* and *y*, for parts of the crystal over ½ mm. distant from the boundary, indicate larger changes than the point *t* on the boundary.

At "20 %" elongation the same characteristics appear, but in Crystals "A" and "B" the position is complicated by the commencement of a shear throughout the crystal, on a second set of planes, as described above. Thus, the points *a* and *m* in Figs. 7 and 8 have probably reached their positions in those diagrams by devious routes and therefore they represent greater changes than

Digitized by Google

the point *e* and the points *e* and *u* respectively. In Crystal "C," Fig. 9, double slipping had not commenced at the "20 °/₀" elongation and there is a definite increase in the change of orientation in passing from the point *u*, on the boundary, through the point *w*, a millimetre away, to the point *x*, in the body of the crystal.

The irregular nature of the surface contours in the vicinity of a crystal boundary introduces a source of uncertainty in the determination of the region which is being subjected to analysis. The boundaries of this region, even on a regular surface, also alter

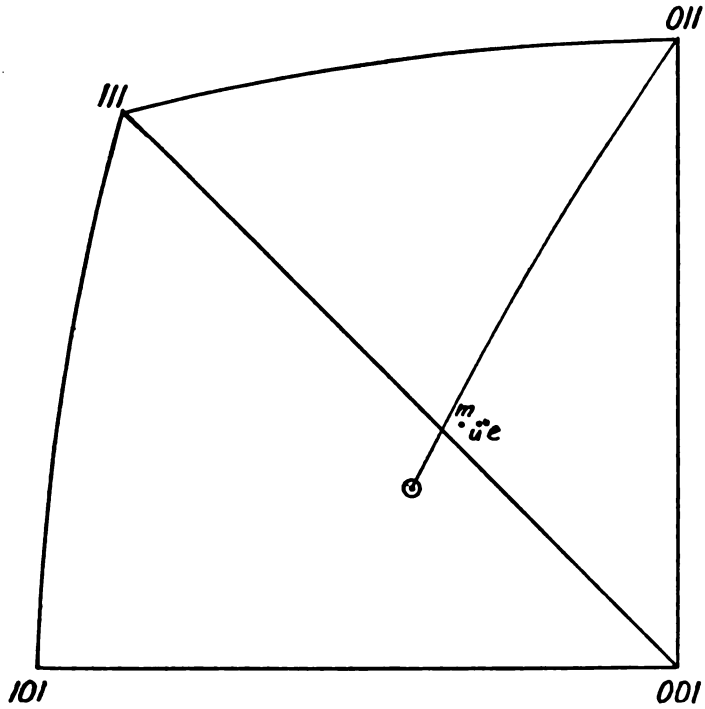


Fig. 8. Crystal "B." Extension "20 °/₀."

slightly with the setting-angle at which a reflection is obtained. The positions indicated by the circles in Figs. 2 and 6 cannot therefore be more than approximate estimates of the regions subjected to analysis. When the orientation is altering rapidly, as it does near to a boundary, it is important not only to know exactly the centre of the region under analysis, but also to confine the region to as small an area as possible. The width of the X-ray beam was a third of that previously used and was estimated to have a diameter of 0·4 mm. at the point of reflection. Measure-

Digitized by Google

ments confined to a smaller distance from the boundary could be made by moving the beam so that the bulk of it was incident upon the adjacent crystal, the reflections from which would in general be in entirely different directions. But it was usually impossible to go very far without losing trace of one of the two axes necessary for a complete analysis, owing to the shifting of the region of incidence of the primary beam as the setting-angle altered.

When a single-crystal is deformed, the X-ray reflections show

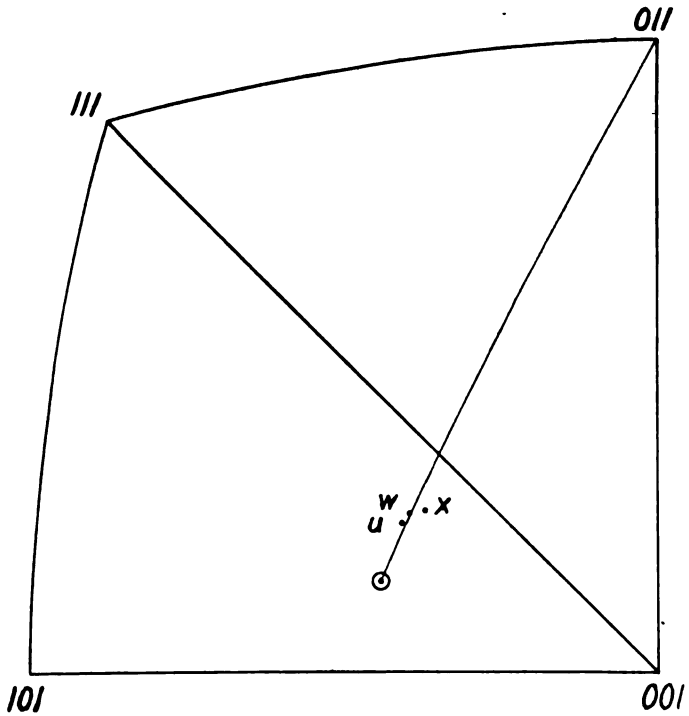


Fig. 9. Crystal "C." Extension "20 %."

a slight breaking-up of the regularity of the crystal structure, which increases as the deformation increases. If the deformation near to a crystal boundary is less than in the body of the crystal, then the trace of the reflected ray on the photographic plate should not show as much "spreading" when from the neighbourhood of a crystal boundary as when from the body of a crystal. But the large change in orientation which may occur in the width of the X-ray beam when in the vicinity of a boundary has a similar effect of spreading the reflected beam. The two effects become confused

Digitized by Google

and it was usually found that very near to a boundary the latter predominated, so that the spots on the photographic became more, instead of less, diffuse. This is illustrated in Figs. 10 to 15* (Plate IV).

Fig. 10 shows the traces obtained from a reflection corresponding to an axis of Crystal "A," before extension. On the left is a line of spots, each of which marks the trace of one half of the undeflected primary beam for the setting-angle denoted by the adjacent number. There is a strong $K\alpha$ reflection for a setting-angle of 38°, a fainter one for 37°, and a $K\beta$ reflection for 34°. Reflections from the continuous spectrum also appear faintly. In Fig. 11 are reflections from the same axis obtained at the point *d* (Fig. 2) after "10 %" elongation. The same reflection, but obtained from the point *f* at "10 %" elongation, is shown in Fig. 12, where the spots are seen to have become more definite than in Fig. 11. (A reflection from an axis of Crystal "B" also appears in Fig. 12.)

Again, Fig. 13 shows a reflection from Crystal "B," before extension, for a setting-angle of 39° and another for 21°. Figs. 14 and 15 show reflections corresponding to the former, after "10 %" elongation, obtained from points *k* and *g* (Fig. 2) respectively. The "spreading" of the reflected beam at a point near to the boundary is again evident.

The two causes of diffuseness in the traces of the reflected ray can only be more effectively separated by further diminishing the width of the beam and thereby considerably increasing the difficulty of the work. The rate of diminution in the change of orientation increases as the boundary is approached and it is possible that at the boundary the deformation is very small indeed.

It is noteworthy that in the two adjacent crystals, "B" and "C," the stereographic diagrams show a departure from a direct approach to the [011] axis which is independent of the distance from the boundary. In Figs. 4 and 8, for Crystal "B," this effect seems greater than could be accounted for by any possible error in the analysis; but in Crystal "A," which occupied the whole of the cross-section of the test-piece, it does not appear. It may be interpreted as a small general shear in the crystal in a direction other than that of the normal shear. In the body of these two crystals there were also found sharply-defined zones in which the deformation was proved to be such a shear, on planes and in directions other than normal. This effect is separate from (although related to) the lessening of the normal deformation as a boundary

* The original plates from which the author intended Figs. 10 to 15 to be reproduced were broken in transit from Australia. The figures here shown were reproduced from prints in which the spots were far less well defined than in the originals.

Digitized by Google

is approached, which is the subject of this paper. It has been studied in further extensions of the same specimen and in other specimens and it is hoped shortly to publish an account of it.

The author wishes to express his thanks to Professor Sir Ernest Rutherford and to Professor G. I. Taylor for their interest in the work and also to the latter for assistance in publishing this paper; among others, especially to Miss Elam, who manufactured the specimens and gave the benefit of her experience in X-ray analysis.

Digitized by Google

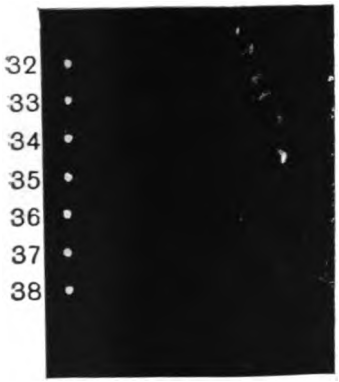


Fig. 10

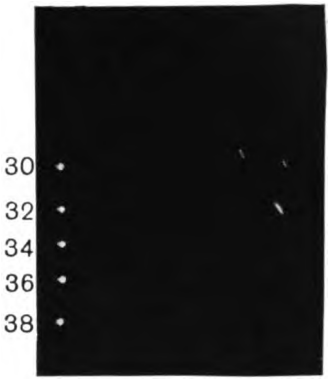


Fig. 11

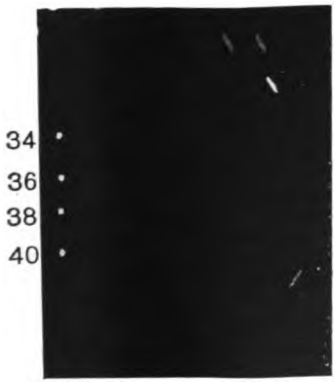


Fig. 12

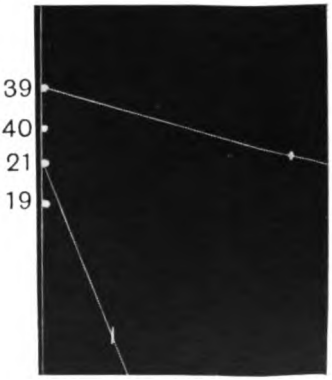


Fig. 13

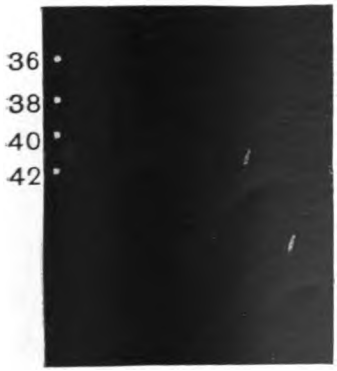


Fig. 14

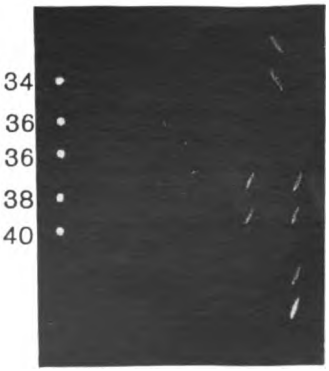


Fig. 15

Digitized by Google

Digitized by Google

Some Developments in the X-ray Analysis of Single-Crystals. By R. L. ASTON, Barker Graduate Scholar of the University of Sydney, Dominion Exhibitioner of Trinity College. (Communicated by Mr G. I. TAYLOR.)

[*Received* 11 November 1926, *read* 19 January 1925.]

A method has been devised* by Dr Alex Müller for determining the orientation of a single-crystal of metal by photographic measurement of the reflection of characteristic X-rays from surface

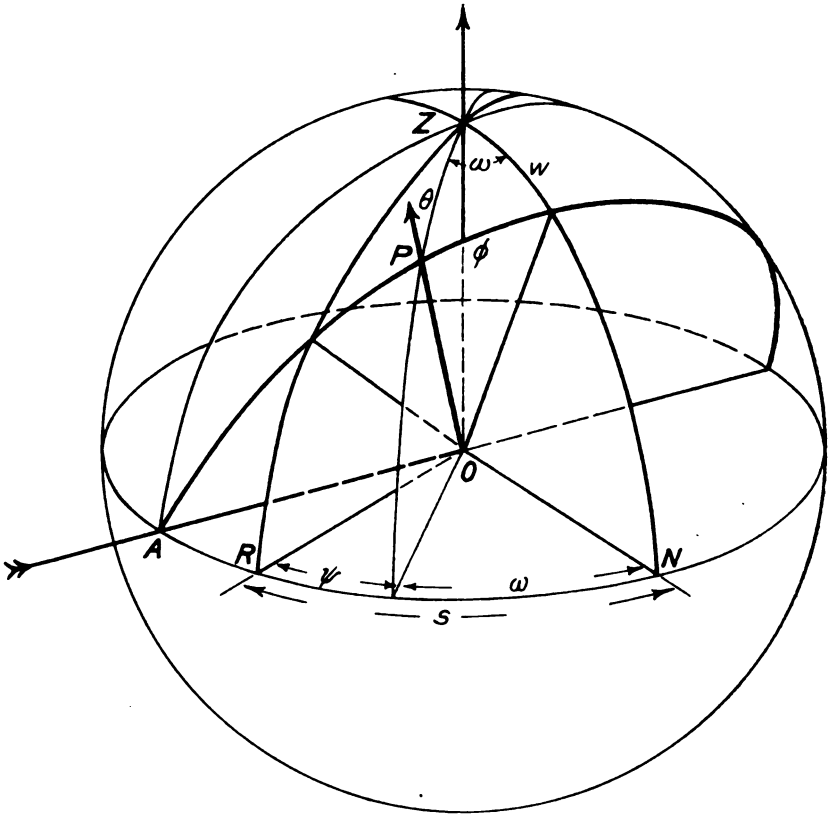


Fig. 1.

layers. The incident beam passes perpendicularly through an axis of rotation around which the crystal is turned until a reflection is obtained with one of the component wave-lengths of the X-rays.

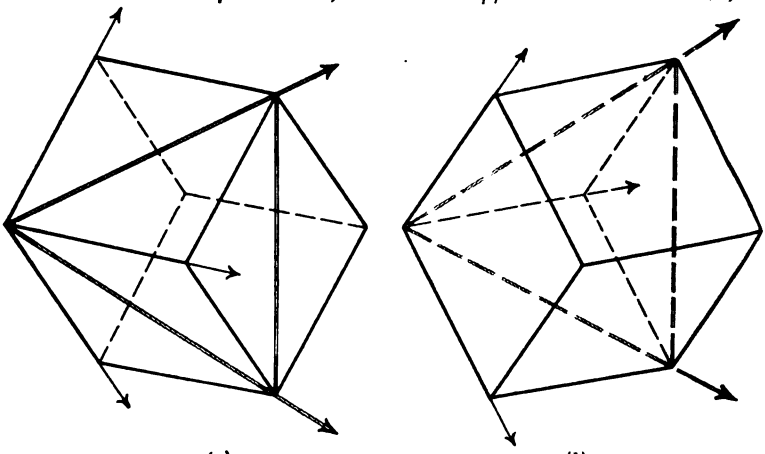
* *Roy. Soc. Proc.* A, **105** (1924), p. 500.

Digitized by Google

In Fig. 1 let AO, OP and OZ be respectively the directions of the primary beam, of the normal to the reflecting planes and of the axis of rotation. The crystal specimen is on the left-hand side of the primary beam. Let ZON be a plane perpendicular to AO and let ZOR be a reference plane chosen in the specimen. For each reflection obtained, values are determined for the reflection angle, ϕ, for the "setting-angle," s, and for the angle w of Fig. 1. From these values the spherical polar co-ordinates (θ, ψ) of the crystal axis concerned may be calculated by the equations

$$\cos\theta = \cos w \,.\, \cos\phi \quad \ldots\ldots\ldots\ldots\ldots\ldots (1),$$

$$\psi = s - \omega\,;\ \sin\omega = \sin\phi/\sin\theta \quad \ldots\ldots\ldots\ldots\ldots (2).$$



(*a*) (*b*)

Fig. 2.

Iron Crystals.

With aluminium crystals Müller used Copper "K" radiation, obtaining reflections from (100) and from (111) planes. Any two such reflections are sufficient to determine the crystal completely.

The method has now been used also for iron crystals. Copper radiation is heavily absorbed in iron and an iron anticathode was substituted, giving reflections only from (110) planes. When depending solely upon [110] type axes, two axes are not in general sufficient to determine the orientation of the crystal. For if the axes determined lie all in the same (111) plane there is an ambiguous solution, as is illustrated by (*a*) and (*b*), Fig. 2, and another axis must be found which does not lie in this plane. It may thus be necessary to have found four of the six [110] axes before the solution is completed.

To facilitate calculations, Tables I and II have been prepared for aluminium and iron crystals with an iron anticathode. They

Digitized by Google

TABLE I. *Values of θ. Iron Anticathode.*

w	ALUMINIUM (100) Kα	ALUMINIUM (100) Kβ	ALUMINIUM (111) Kα	ALUMINIUM (111) Kβ	IRON (110) Kα	IRON (110) Kβ	w	ALUMINIUM (100) Kα	ALUMINIUM (100) Kβ	ALUMINIUM (111) Kα	ALUMINIUM (111) Kβ	IRON (110) Kα	IRON (110) Kβ
°	°	°	°	°	°	°	°	°	°	°	°	°	°
0	28·3	25·5	24·3	21·9	28·9	26·0	45	51·5	50·3	49·9	49·0	51·7	50·5
1	28·4	25·6	24·3	21·9	28·9	26·0	46	52·3	51·2	50·7	49·9	52·5	51·4
2	28·4	25·6	24·4	22·0	29·0	26·1	47	53·1	52·0	51·6	50·7	53·3	52·2
3	28·5	25·7	24·4	22·1	29·0	26·2	48	53·9	52·9	52·4	51·6	54·1	53·0
4	28·6	25·8	24·6	22·3	29·1	26·3	49	54·7	53·7	53·3	52·5	54·9	53·9
5	28·8	26·0	24·8	22·5	29·3	26.4	50	55·5	54·5	54·1	53·4	55·7	54·7
6	28·9	26·2	25·0	22·7	29·5	26·6	51	56·4	55·4	55·0	54·3	56·6	55·6
7	29·1	26·4	25·2	23·0	29·6	26·9	52	57·2	56·3	55·9	55·2	57·4	56·4
8	29·4	26·7	25·5	23·3	29·9	27·1	53	58·0	57·1	56·7	56·1	58·2	57·3
9	29·7	27·0	25·8	23·6	30·1	27·4	54	58·9	58·0	57·6	57·0	59·0	58·1
10	29·9	27·3	26·1	24·0	30·4	27·7	55	59·7	58·8	58·5	57·9	59·8	59·0
11	30·3	27·7	26·5	24·4	30·7	28·1	56	60·5	59·7	59·4	58·8	60·6	59·8
12	30·6	28·0	26·9	24·8	31·1	28·5	57	61·4	60·6	60·2	59·7	61·5	60·7
13	31·0	28·4	27·4	25·3	31·4	28·9	58	62·2	61·4	61·1	60·6	62·4	61·6
14	31·4	28·9	27·8	25·8	31·8	29·3	59	63·1	62·3	62·0	61·5	63·2	62·4
15	31·8	29·4	28·3	26·3	32·2	29·7	60	63·9	63·2	62·9	62·4	64·0	63·3
16	32·2	29·8	28·8	26·9	32·7	30·2	61	64·7	64·1	63·8	63·3	64·9	64·2
17	32·7	30·3	29·3	27·5	33·1	30·7	62	65·6	64·9	64·7	64·2	65·6	65·0
18	33·2	30·9	29·9	28·1	33·6	31·3	63	66·5	65·8	65·6	65·1	66·6	65·9
19	33·7	31·4	30·5	38·7	34·1	31·8	64	67·3	66·7	66·4	66·0	67·4	66·8
20	34·2	32·0	31·1	29·3	34·6	32·4	65	68·2	67·6	67·3	66·9	68·3	67·7
21	34·7	32·6	31·7	30·0	35·2	33·0	66	69·0	68·5	68·2	67·8	69·1	68·6
22	35·3	33·2	32·3	30·7	35·7	33·6	67	69·9	69·4	69·1	68·7	70·0	69·4
23	35·9	33·8	33·0	31·4	36·3	34·2	68	70·8	70·2	70·0	69·7	70·9	70·3
24	36·5	34·5	33·6	32·1	36·9	34·8	69	71·6	71·1	70·9	70·6	71·7	71·2
25	37·1	35·1	34·3	32·8	37·5	35·4	70	72·5	72·0	71·8	71·5	72·6	72·1
26	37·7	35·8	35·0	33·5	38·1	36·1	71	73·4	72·9	72·7	72·4	73·4	73·0
27	38·4	36·5	35·7	34·2	38·7	36·8	72	74·2	73·8	73·6	73·3	74·3	73·9
28	39·0	37·2	36·4	35·0	39·4	37·5	73	75·1	74·7	74·3	74·3	75·2	74·8
29	39·7	37·9	37·1	35·8	40·0	38·2	74	76·0	75·6	75·4	75·2	76·0	75·7
30	40·4	38·6	37·9	36·5	40·7	38·9	75	76·8	76·5	76·4	76·1	76·9	76·5
31	41·0	39·3	38·6	37·3	41·4	39·6	76	77·7	77·4	77·3	77·0	77·8	77·4
32	41·7	40·1	39·4	38·1	42·1	40·3	77	78·6	78·3	78·2	78·0	78·6	78·3
33	42·4	40·8	40·1	38·9	42·7	41·1	78	79·5	79·2	79·1	78·9	79·5	79·2
34	43·1	41·6	40·9	39·7	43·5	41·8	79	80·3	80·1	80·0	79·8	80·4	80·1
35	43·9	42·3	41·7	40·5	44·2	42·6	80	81·2	81·0	80·9	80·7	81·3	81·0
36	44·6	43·1	42·5	41·4	44·9	43·3	81	82·1	81·9	81·8	81·7	82·1	81·9
37	45·4	43·9	43·3	42·2	45·6	44·1	82	83·0	82·8	82·7	82·6	83·0	82·8
38	46·1	44·7	44·1	43·0	46·4	44·9	83	83·8	83·7	83·6	83·5	83·9	83·7
39	46·8	45·5	44·9	43·9	47·1	45·7	84	84·7	84·6	84·5	84·4	84·7	84·6
40	47·6	46·3	45·7	44·7	47·9	46·5	85	85·6	85·5	85·4	85·4	85·6	85·5
41	48·4	47·1	46·5	45·6	48·6	47·3	86	86·5	86·4	86·3	86·3	86·4	86·4
42	49·2	47·9	47·4	46·4	49·4	48·1	87	87·4	87·3	87·3	87·2	87·4	87·3
43	49·9	48·7	48·2	47·3	50·2	48·9	88	88·2	88·2	88·2	88·1	88·2	88·2
44	50·3	49·5	49·0	48·1	51·0	49·7	89	89·1	89·1	89·1	89·1	89·1	89·1
45	51·5	50·3	49·9	49·0	51·7	50·5	90	90·0	90·0	90·0	90·0	90·0	90·0

Digitized by Google

TABLE II. *Values of ω. Iron Anticathode.*

w	ALUMINIUM (100)		ALUMINIUM (111)		IRON (110)		w	ALUMINIUM (100)		ALUMINIUM (111)		IRON (110)	
	Kα	Kβ	Kα	Kβ	Kα	Kβ		Kα	Kβ	Kα	Kβ	Kα	Kβ
°	°	°	°	°	°	°	°	°	°	°	°	°	
0	90·0	90·0	90·0	90·0	90·0	90·0	45	37·3	34·0	32·5	29·6	37·9	34·
1	88·1	87·9	87·7	87·5	88·1	87·9	46	36·9	33·6	32·1	29·2	37·5	34·
2	86·3	85·9	85·6	85·1	86·4	86·0	47	36·4	33·1	31·7	28·8	37·0	33·
3	84·3	83·7	83·4	82·7	84·4	83·9	48	36·0	32·7	31·3	28·4	36·6	33·
4	82·6	81·6	81·1	80·2	82·7	81·9	49	35·6	32·3	30·9	28·1	36·2	32·
5	80·8	79·5	79·0	77·8	81·1	79·9	50	35·2	31·9	30·5	27·7	35·8	32·
6	79·0	77·6	77·0	75·5	79·3	77·9	51	34·8	31·6	30·1	27·4	35·4	32·
7	77·3	75·7	75·0	73·2	77·5	76·0	52	34·4	31·2	29·8	27·1	35·0	31·
8	75·6	73·8	73·0	70·9	75·8	74·1	53	34·0	30·9	29·5	26·7	34·6	31·
9	73·9	71·9	70·9	68·8	74·2	72·2	54	33·7	30·5	29·1	26·4	34·3	31·
10	72·2	70·0	68·9	66·7	72·6	70·4	55	33·4	30·2	28·8	26·2	33·9	30·
11	70·5	68·2	67·1	64·6	70·9	68·6	56	33·1	29·9	28·5	25·9	33·6	30·
12	68·9	66·5	65·3	62·7	69·3	66·9	57	32·8	29·7	28·3	25·6	33·3	30·
13	67·3	64·8	63·5	60·8	67·8	65·2	58	32·5	29·4	28·0	25·4	33·1	29·
14	65·8	63·1	61·8	59·0	66·3	63·6	59	32·2	29·1	27·8	25·1	32·8	29·
15	64·3	61·5	60·2	57·3	64·9	62·0	60	31·9	28·9	27·5	24·9	32·5	29·
16	62·9	60·0	58·6	55·6	63·5	59·5	61	31·7	28·6	27·3	24·7	32·2	29·
17	61·5	58·5	57·1	54·0	62·3	59·1	62	31·4	28·4	27·1	24·5	32·0	28·
18	60·2	57·1	55·6	52·5	60·9	57·7	63	31·2	28·2	26·9	24·3	31·7	28·
19	58·9	55·7	54·2	51·0	59·5	56·3	64	31·0	28·0	26·7	24·1	31·3	28·
20	57·6	54·4	52·9	49·6	58·2	55·0	65	30·8	27·8	26·5	23·9	31·3	28·
21	56·4	53·1	51·6	48·3	57·0	53·7	66	30·6	27·6	26·3	23·8	31·1	28·
22	55·2	51·9	50·3	47·1	55·8	52·5	67	30·4	27·4	26·1	23·6	30·9	27·
23	54·1	50·7	49·1	45·9	54·7	51·3	68	30·2	27·2	25·9	23·5	30·8	27·
24	53·0	49·6	48·0	44·7	53·6	50·2	69	30·0	27·1	25·8	23·3	30·6	27·
25	51·9	48·5	46·9	43·6	52·6	49·1	70	29·9	26·9	25·6	23·2	30·4	27·
26	50·9	47·5	45·8	42·5	51·5	48·0	71	29·7	26·8	25·5	23·1	30·2	27·
27	49·9	46·5	44·8	41·5	50·5	47·0	72	29·6	26·7	25·4	22·9	30·1	27·
28	49·0	45·5	43·8	40·6	49·6	46·1	73	29·4	26·5	25·3	22·8	29·9	27·
29	48·1	44·6	42·9	39·7	48·7	45·2	74	29·3	26·4	25·1	22·7	29·8	26·
30	47·2	43·7	42·0	38·8	47·8	44·3	75	29·2	26·3	25·0	22·6	29·7	26·
31	46·3	42·8	41·2	38·0	47·0	43·4	76	29·1	26·2	24·9	22·5	29·6	26·
32	45·5	42·0	40·4	37·2	46·2	42·6	77	29·0	26·1	24·8	22·4	29·5	26·
33	44·7	41·2	39·6	36·5	45·4	41·8	78	28·9	26·0	24·8	22·4	29·4	26·
34	44·0	40·5	38·9	35·7	44·6	41·1	79	28·8	25·9	24·7	22·3	29·3	26·
35	43·3	39·8	38·2	35·0	43·9	40·4	80	28·7	25·9	24·6	22·2	29·3	26·
36	42·6	39·1	37·5	34·4	43·2	39·7	81	28·6	25·8	24·6	22·2	29·2	26·
37	41·9	38·4	36·8	33·8	42·5	39·0	82	28·6	25·7	24·5	22·1	29·1	26·
38	41·2	37·8	36·2	33·2	41·9	38·4	83	28·5	25·7	24·5	22·1	29·1	26·
39	40·6	37·2	35·6	32·6	41·2	37·8	84	28·5	25·6	24·4	22·0	29·0	26·
40	40·0	36·6	35·1	32·0	40·6	37·2	85	28·4	25·6	24·4	22·0	29·0	26·
41	39·4	36·1	34·5	31·5	40·0	36·6	86	28·4	25·6	24·3	22·0	28·9	26·
42	38·9	35·5	34·0	31·0	39·5	36·1	87	28·4	25·6	24·3	21·9	28·9	26·
43	38·3	35·0	33·5	30·5	39·0	35·6	88	28·4	25·5	24·3	21·9	28·9	26·
44	37·8	34·5	33·0	30·1	38·5	35·1	89	28·4	25·5	24·3	21·9	28·9	26·
45	37·3	34·0	32·5	29·6	37·9	34·6	90	28·3	25·5	24·3	21·9	28·9	26·

Digitized by Google

TABLE III. *Values of θ.*

ϕ / w	5°	10°	15°	20°	25°	30°	35°
°	°	°	°	°	°	°	°
5	7·1	11·2	15·8	20·6	25·5	30·6	35.3
10	11·2	14·1	17·0	22·3	26·8	31·5	36·2
15	15·8	17·0	21·1	24·8	28·9	33·2	37·7
20	20·6	22·3	24·8	28·0	31·6	35·5	39·7
25	25·5	26·8	28·9	31·6	34·8	38·3	42·1
30	30·4	31·5	33·2	35·5	38·3	41·4	44·8
35	35·3	36·2	37·7	39·7	42·1	44·8	47·8
40	40·3	41·0	42·3	44·0	46·0	48·4	51·1
45	45·2	45·9	46·9	48·4	50·1	52·2	54·6
50	50·2	50·7	51·6	52·8	54·4	56·2	58·2
55	55·2	55·6	56·4	57·4	58·7	60·2	62·0
60	60·1	60·5	61·1	62·0	63·2	64·3	65·8
65	65·1	65·4	65·9	66·6	67·5	68·5	69·7
70	70·1	70·3	70·7	71·3	71·9	72·8	73·7
75	75·1	75·2	75·5	75·9	76·4	77·0	77·8
80	80·0	80·2	80·3	80·6	80·9	81·4	81·8
85	85·0	85·1	85·2	85·3	85·5	85·7	85·9
90	90·0	90·0	90·0	90·0	90·0	90·0	90·0

TABLE IV. *Values of ω.*

ϕ / w	5°	10°	15°	20°	25°	30°	35°
°	°	°	°	°	°	°	°
5	45·0	63·7	72·0	76·5	79·4	81·4	82·9
10	26·7	45·4	57·0	64·5	69·6	73·3	76·1
15	18·7	34·3	46·0	54·6	61·0	65·9	69·7
20	14·3	27·3	38·1	46·8	53·7	59·4	64·0
25	11·7	22·7	32·4	40·7	47·8	53·8	58·9
30	9·9	19·4	28·2	36·1	43·0	49·1	54·5
35	8·7	17·1	25·0	32·4	39·1	45·2	50·7
40	7·8	15·3	22·6	29·5	36·0	41·9	47·5
45	7·1	14·0	20·8	27·2	33·4	39·2	44·7
50	6·5	13·0	19·3	25·4	31·3	37·0	42·4
55	6·1	12·2	18·1	24·0	29·6	35·2	40·5
60	5·8	11·5	17·2	22·8	28·3	33·7	39·0
65	5·5	11·0	16·5	21·9	27·2	32·5	37·7
70	5·3	10·6	15·9	21·2	26·4	31·6	36·7
75	5·2	10·3	15·5	20·6	25·8	30·9	35·9
80	5·1	10·2	15·2	20·3	25·3	30·4	35·4
85	5·0	10·0	15·1	20·1	25·1	30·1	35·1
90	5·0	10·0	15·0	20·0	25·0	30·0	35·0

Digitized by Google

give values of θ and of ω for intervals in w of one degree. The values taken for the side of unit cube are $4{\cdot}07 \times 10^{-8}$ cm. for aluminium* and $2{\cdot}83 \times 10^{-8}$ cm. for iron†, while the wave-lengths of the $K\alpha$ and $K\beta$ rays are taken as $1{\cdot}93239 \times 10^{-8}$ cm. and $1{\cdot}75397 \times 10^{-8}$ cm. respectively.

Use of the Continuous Spectrum.

It will be shown below that, at a given point on the surface of a specimen, the position of an axis can be found only if it lies within a certain zone, whose boundaries vary with the reflection angle. Thus it is often difficult or impossible to determine axes sufficient for a complete analysis with the characteristic rays only of the anticathode metal selected. It is then advantageous or necessary to use radiation of another wave-length, thereby changing to an accessible value the reflection angles corresponding to other axes. Instead of substituting another metal for this purpose, one may use the continuous spectrum of the same anticathode. Thus, with an aluminium crystal, an iron anticathode in a Shearer tube at 50,000 volts gave reflections from the shorter wave-lengths of the continuous spectrum which were comparable in intensity with those from the characteristic rays.

As the range of wave-lengths available is both wide and continuous, reflections usually appear from one or more sets of planes at any setting-angle tried, especially for small angles of incidence to the surface. This contrasts with the time and trouble often spent in searching for the proper setting-angle, if any, to obtain a reflection from the characteristic rays. There is a slight decrease in accuracy, which is largely compensated for by the readiness with which measurements may be made at several angles for each axis. Moreover the wave-lengths of the rays reflected and the types of the reflecting planes are not known and are not immediately determinable. But the position of an axis may be calculated from the corresponding reflection angles, which may be measured directly for each reflection. From the values obtained for ϕ, w and s, the co-ordinates may be found by interpolation in Tables III and IV, which give values of θ and of ω at intervals of 5° in ϕ and w.

When several axes have been determined, the angle between each pair of them may be calculated and the values found may be compared with those which exist between different types of axes of the system to which the crystal belongs. Iron and aluminium are both of the cubic system and strong reflections are obtained from the (100), (110), (111) and (211) types of planes. The pairs

* Scherrer, *Phys. Zeitschr.* **19** (1918), p. 23.
† Westgren, "Engineering," **111** (1921), p. 727.

Digitized by Google

of axes which correspond to each possible angle in this system are given in Table V. From such a comparison the nature of each axis may be determined without ambiguity. Usually three or four axes are sufficient for the purpose.

It will be seen that whereas the labour of exploring for axes is considerably diminished by the use of the continuous spectrum, the amount of subsequent measurement and calculation is greater than when an analysis can be made with the characteristic spectrum. A complete and accurate analysis may always be made with the continuous spectrum only. A few exposures, however, will usually give reflections from the continuous spectrum sufficient to calculate the positions of several axes approximately and thence to determine the conditions, if any, under which reflections from

TABLE V.

Angle	Pairs of Axes
°	
90·0	100/100, 100/110, 110/110, 110/111, 110/211, 111/211
80·4	211/211
73·2	110/211
70·5	111/111, 211/211
65·9	100/211
61·9	111/211
60·0	110/110, 211/211
54·7	100/111, 110/211
48·2	211/211
45·0	100/110
35·3	100/211, 110/111
33·6	211/211
30·0	110/211
19·5	111/211

the characteristic spectrum might be expected. Again, in the case of the ambiguous solution referred to above, a continuous spectrum reflection may readily be obtained which will give another axis with sufficient accuracy to discriminate between the two solutions.

Figs. 3 and 4 illustrate the photographic impressions of some of the reflections obtained from an aluminium crystal with the continuous spectrum of iron. The photographic plate is perpendicular to the primary beam. Exposures were made at several setting-angles with each plate and each semi-circular spot marks the trace of one-half of the primary beam for the setting-angle denoted by the adjacent number, the other half being obstructed by the crystal specimen. The line given by these spots is parallel to the axis of rotation, *OZ* Fig. 1. Fig. 3 shows a series of reflections corresponding to one crystal axis at intervals of 2° in the

Digitized by Google

setting-angle. Fig. 4 shows a series of exposures at intervals of one degree in the setting-angle, in which reflections appear at the same time both from the characteristic spectrum and from the continuous spectrum. The two types of reflections may immediately be distinguished by the shapes of the spots, those from the continuous spectrum being elongated in the direction of the line from the corresponding "zero" spots, while those from the monochromatic spectrum are elongated in a perpendicular direction. As the setting-angle is increased the reflection angle for the axis increases and a longer wave-length is selected from the continuous spectrum. Thus the "continuous" reflection is seen first to cross

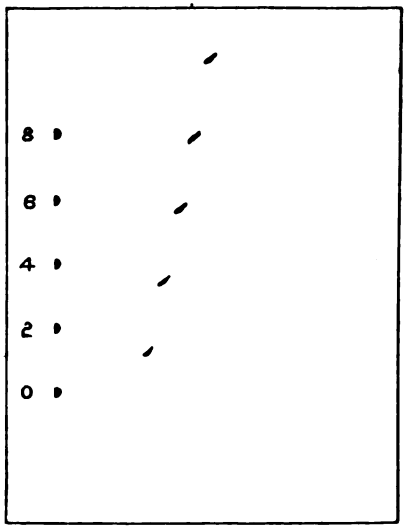


Fig. 3.

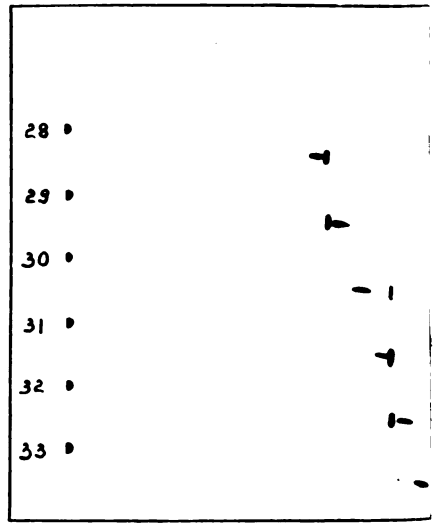


Fig. 4.

the "$K\beta$" reflection and then the "$K\alpha$" reflection. Observation of the crossing of the two series of spots is of assistance in making an accurate estimate of the setting-angle proper to the "characteristic" reflection.

Accessibility of Axes.

The method of analysis gives information only as to the surface of the specimen. It has the further limitation that a crystal axis can be determined only if it lies in such a position that the incident and emergent beams involved in the analysis are both unobstructed by the body of the crystal. For a given reflection angle there is thus defined a cone within which the

Digitized by Google

direction of an axis must lie if it is to be detected in an analysis of a given region of the surface. If the surface is regular in the region under analysis, the limits are given approximately by a grazing incidence or emergence of the rays. The reference plane may usually be conveniently chosen normal to the surface. The term "setting-angle" will be taken hereinafter to mean the angle between the incident beam and the surface. There is thus imposed on the setting-angle a lower limit 0° and an upper limit which varies with the inclination of the crystal axis concerned.

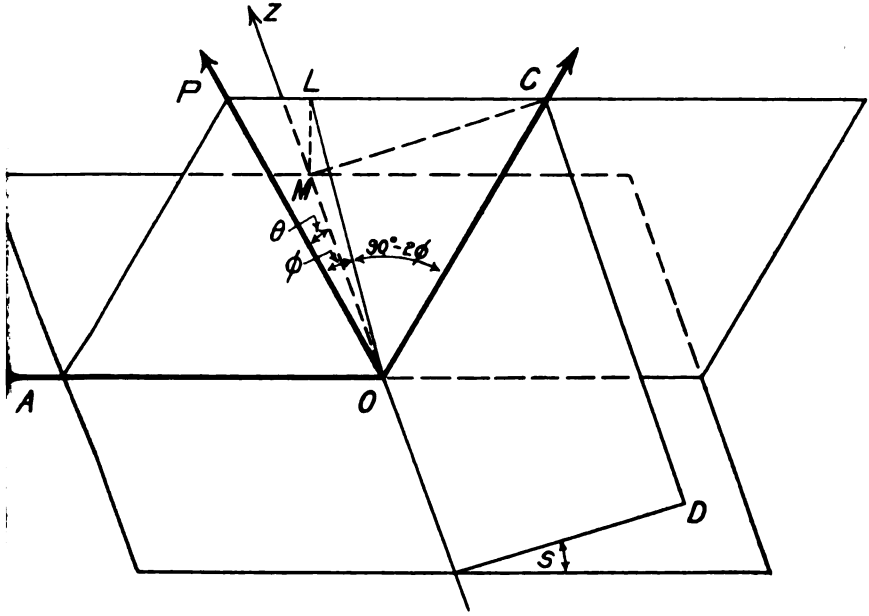


Fig. 5.

Fig. 5 illustrates the geometry of the latter case. AO, OC and OP are respectively the directions of the incident and emergent beams and of the crystal axis; OZ the axis of rotation; and DOZ the plane tangential to the surface, which has been turned until OC is grazing it. The angle between the planes ZOA and ZOD will measure the setting-angle s.

Let OL be drawn perpendicular to AO in the plane AOC, and let CM and LM be perpendicular to OZ. From the solid angle (O, PML)

$$\cos\theta = \cos\phi \,.\, \cos \angle MOL.$$

Digitized by Google

Then $$\tan s = \frac{ML}{CL} = \frac{OL \,.\, \sin \angle MOL}{OL \,.\, \cot 2\phi}$$

$$= \sqrt{1 - \frac{\cos^2 \theta}{\cos^2 \phi}} \Big/ \cot 2\phi.$$

Thus $$\tan s = \frac{2 \sin \phi}{\cos 2\phi} \sqrt{\cos^2 \phi - \cos^2 \theta}.$$

TABLE VI.

Radiation	Iron $K\alpha$								
Crystal	IRON			ALUMINIUM					
Plane	(110)			(100)			(111)		
ϕ	28°·9			28°·3			24°·3		
Limit	Lower	Upper		Lower	Upper		Lower	Upper	
θ	ψ	s	ψ	ψ	s	ψ	ψ	s	ψ
	°	°	°	°	°	°	°	°	°
25	—	—	—	—	—	—	−76·6	6·9	−69·7
30	−74·9	13·2	−61·7	−71·6	15·7	−55·9	−53·7	19·5	−34·2
35	−57·3	29·2	−18·1	−55·8	29·1	−26·7	−45·8	26·4	−19·4
40	−48·7	37·5	−11·2	−47·7	36·8	−10·9	−39·8	31·6	−18·2
45	−43·1	43·1	0·0	−42·3	42·3	0·0	−35·6	35·6	0·0
50	−39·1	47·1	8·0	−38·2	46·1	7·9	−32·5	38·8	6·3
55	−36·1	50·1	14·0	−35·4	49·1	13·7	−30·1	41·4	11·3
60	−33·9	52·4	18·5	−33·3	51·4	18·1	−28·3	43·4	15·1
65	−32·2	54·2	22·0	−31·6	53·2	21·6	−27·0	45·1	18·1
70	−30·9	55·6	24·7	−30·4	54·5	24·1	−25·9	46·4	20·5
75	−30·0	56·5	26·5	−29·4	55·5	26·1	−25·2	47·4	22·2
80	−29·4	57·2	27·8	−28·8	56·2	27·4	−24·7	48·0	23·3
85	−29·0	57·6	28·6	−28·5	56·6	28·1	−24·4	48·4	24·0
90	−28·9	57·7	28·9	−28·3	56·7	28·3	−24·3	48·6	24·3

This equation, with the equations (2) above, defines the co-ordinates (θ, ψ) of one of the boundaries of the region of accessibility. The other boundary is given by the equations (2) with $s = 0$. Table VI gives the limiting values of s and of ψ for the reflection angles which occur with iron or aluminium crystals under iron $K\alpha$ radiation. Table VII gives the values for reflection angles up to 35° at intervals of 5°. The boundaries of the region of accessibility are shown in Fig. 6 plotted in stereographic projection for reflection angles of 10°, 20°, and 30°. The pole of the projection is the axis of rotation (OZ, Fig. 1) and the arrow denotes the direction of the normal to the surface in the region

Digitized by Google

TABLE VII.

Values of ψ at Lower Limit.

ϕ / θ	5°	10°	15°	20°	25°	30°	35°
°	°						
5	−90·0	°					
10	−30·1	−90·0	°				
15	−19·7	−42·1	−90·0	°			
20	−14·8	−30·5	−49·2	−90·0	°		
25	−11·9	−24·3	−37·8	−54·0	−90·0	°	
30	−10·0	−20·3	−31·2	−43·2	−57·7	−90·0	°
35	− 8·7	−17·6	−26·8	−36·6	−47·5	−60·7	−90·0
40	− 7·8	−15·7	−23·7	−32·1	−41·1	−51·1	−63·2
45	− 7·1	−14·2	−21·5	−28·9	−36·7	−45·0	−54·2
50	− 6·5	−13·1	−19·7	−26·5	−33·5	−40·7	−48·5
55	− 6·1	−12·2	−18·4	−24·7	−31·1	−37·6	−44·4
60	− 5·8	−11·6	−17·4	−23·3	−29·2	−35·3	−41·5
65	− 5·5	−11·0	−16·6	−22·2	−27·8	−33·5	−39·3
70	− 5·3	−10·6	−16·0	−21·3	−26·7	−32·1	−37·6
75	− 5·2	−10·4	−15·5	−20·7	−25·9	−31·2	−36·4
80	− 5·1	−10·2	−15·2	−20·3	−25·4	−30·5	−35·6
85	− 5·0	−10·0	−15·1	−20·1	−25·1	−30·1	−35·2
90	− 5·0	−10·0	−15·0	−20·0	−25·0	−30·0	−35·0

Lower limit of $s = 0$.

Values of ψ at Upper Limit.

ϕ / θ	5°	10°	15°	20°	25°	30°	35°
°	°						
5	−90·0	°					
10	−28·6	−90·0	°				
15	−17·2	−38·0	−90·0	°			
20	−11·4	−24·3	−41·6	−90·0	°		
25	− 7·9	−16·2	−26·3	−41·5	−90·0	°	
30	− 5·0	−10·5	−16·9	−25·2	−28·3	−90·0	°
35	− 3·0	− 6·2	− 9·8	−14·2	−20·5	−31·4	−90·0
40	− 1·4	− 2·8	− 4.3	− 6·2	− 8·6	−12·2	−19·0
45	0·0	0·0	0·0	0·0	0·0	0·0	0·0
50	1·2	2·3	3·6	5·0	6·5	8·6	11·1
55	2·1	4·3	6·5	8·9	11·6	14·8	18·6
60	2·9	5·8	8·7	12·1	15·6	19·4	23·8
65	3·6	7·2	10·8	14·6	18·7	23·0	27·7
70	4·1	8·2	12·4	16·7	21·2	25·8	30·6
75	4·5	9·0	13·6	18·2	22·9	27·6	32·6
80	4·7	9·5	14·4	19·2	24·1	29·0	34·0
85	4·9	9·9	14·8	19·8	24·8	29·8	34·7
90	5·0	10·0	15·0	20·0	25·0	30·0	35·0

Upper limit of s = Difference of upper and lower values of ψ.

Digitized by Google

analysed. It may be noticed that the method will not find an axis for which θ is less than ϕ. The range of ψ is always equal to 2ω and for $\theta = 90°$ is from $-\phi$ to $+\phi$. The greatest setting-angle at which a reflection is possible is 2ϕ. Thus, with the continuous spectrum, for a reflection angle of 10° it is useless to try a setting-angle of more than 20°.

If the diagram of the region accessible for any reflection angle be superimposed upon a stereographic projection of the crystal axes, drawn with the same axes of projection, then the axes which fall within the diagram and are accessible for measurement with that reflection angle may immediately be picked out for any

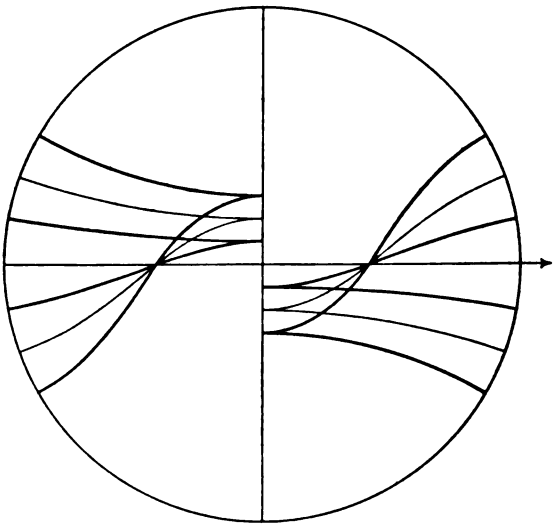


Fig. 6.

portion of the surface. If the direction of the normal to the surface moves in azimuth, the region accessible moves until, when the direction of the normal has been reversed, the whole of the region for which $\theta > \phi$ has been covered. Reversing the specimen in azimuth does not alter the region accessible; but inverting it with regard to the axis of rotation makes a partial alteration, especially for the smaller values of θ. Fig. 7 shows the relation of the accessibility diagram for the specimen inverted to that for the specimen upright, when $\phi = 25°$.

The accessibility diagrams are of use as soon as a solution, or an approximate solution, of the orientation of a crystal has been obtained. It is then possible immediately to determine for which axes reflections may be expected upon any face of the specimen

Digitized by Google

on which the crystal is exposed, or upon which face a reflection may be expected for any given axis. With iron crystals, for example, it was thus possible to discover at once whether sufficient axes could be determined for a complete and unambiguous analysis with the characteristic radiation used.

The case of a rod of circular cross-section is of especial interest. In investigating the deformation of large aluminium crystals in tensile test-pieces, for example, it was necessary to make a series of complete and independent analyses around the periphery of the rod. Fig. 9 illustrates a type of diagram drawn to expedite the work. The axes of the crystal are shown in stereo-

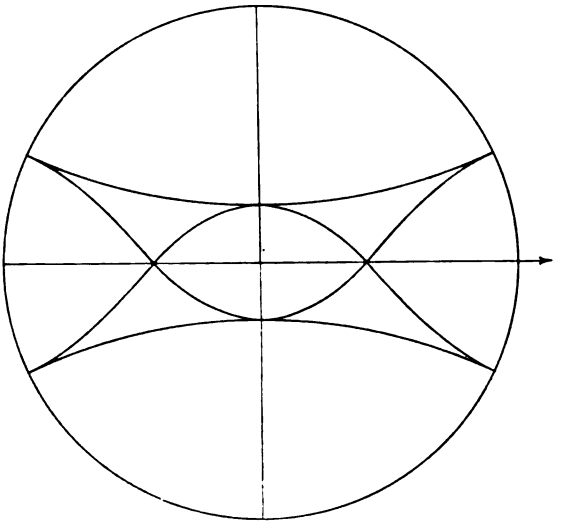


Fig. 7.

graphic projection in Fig. 8, where they have been distinguished by letters in order of increasing inclination to the axis of projection. The large letters denote cubic axes and the small letters triad axes. In Fig. 9 one circle has been allotted to each axis and four reference planes marked on the specimen have been drawn and numbered. Corresponding to the part of the circumference of the rod on which it was possible, provided that the crystal came to the surface in that part, to obtain a reflection for an axis with the characteristic spectrum of iron, a heavy arc has been drawn alongside of the circle representing it in the diagram. An arc drawn on the outside of the circle indicates that a reflection can be found in that part with the rod in the "upright" position, while an arc on the inside indicates the "inverted" position. From

Digitized by Google

such a diagram it is possible to see at a glance which axes may be determined at any point on the circumference, with the specimen either upright or inverted.

Three circular bands have been drawn closer to the centre of the diagram, extending only as far as the crystal is exposed around the periphery of the rod. In parts where it is impossible to determine two axes with the rod in the "upright" position, the outermost of these bands has been blackened. Similarly, the innermost band has been blackened for the "inverted" position. Where it was impossible to determine two axes by combining both positions, the centre circle has also been blackened.

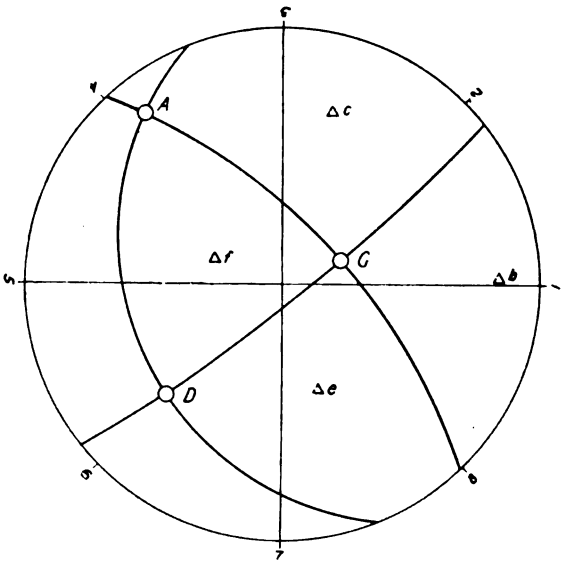


Fig. 8.

The diagram shows that the more steeply inclined axes are accessible over a much smaller range than those nearer to the "horizontal," but that, whereas the range of the latter is in much the same position, whether for the upright or the inverted position of the crystal, when the axis is more steeply inclined, the two ranges overlap very little or may be widely separated. Expressed as a percentage of the length of the circumference in which the crystal was exposed, the sum of the lengths in which it could not be analysed by seeking reflections with the crystal in the "upright" position, in the "inverted" position, or in the two combined, were respectively 62 %, 65 % and 46 %. For six typical crystals the

Digitized by Google

average values of these percentages were respectively 55 °/₀, 54 °/₀ and 25 °/₀. The longest continuous range of the circumference on which an analysis could not be made, or in which it is theoretically possible only for an interval of one or two degrees, was 66° for the upright or inverted positions singly and 50° for a combination of the two positions.

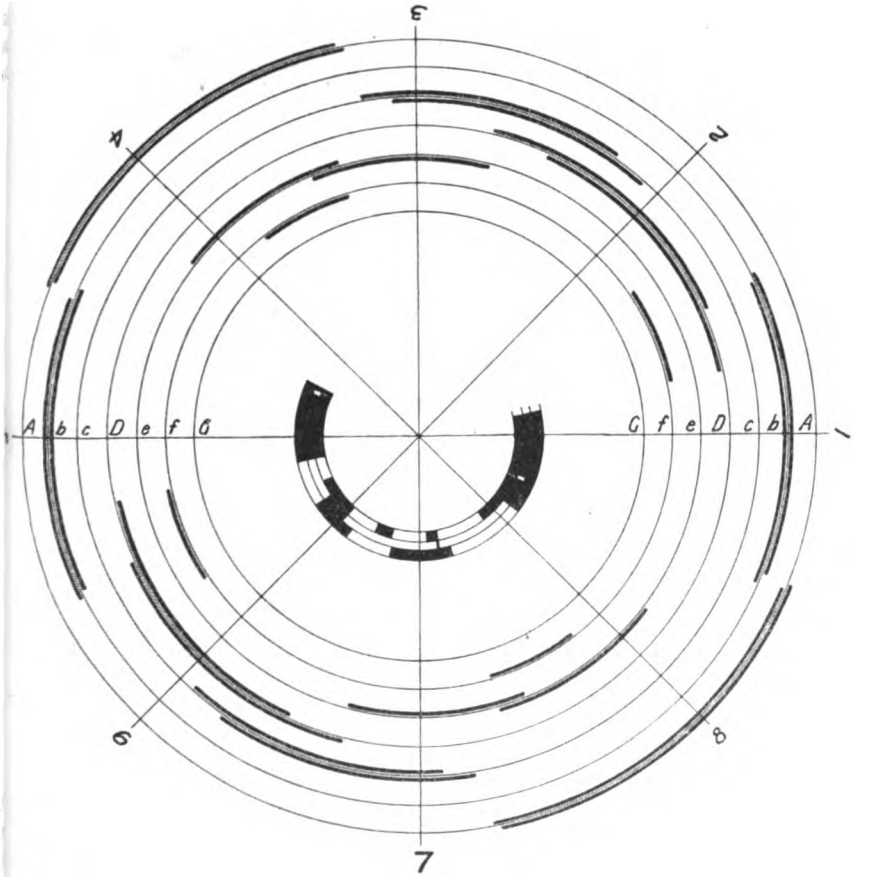


Fig. 9.

To penetrate into these "dark" regions it would be necessary to use radiations of other wave-lengths, to measure other types of axes, or to rotate the crystal about an axis perpendicular to that of the spectrometer, into other orientations. Such a rotation is impossible if the length of the specimen is great compared with the dimensions of the spectrometer, but when the crystal is exposed only on

Digitized by Google

one face, as in a disc, it is the only means of bringing new axes within reach of analysis. In this case the specimen is mounted so that the spectrometer axis lies along the plane face, and is rotated to any orientation about an axis normal to the face. The accessibility diagram is then most conveniently drawn with this latter axis as the axis of projection. Such a projection is drawn for $\phi = 25°$ in Fig. 10, the arrow denoting the "upward" direction (*OZ*, Fig. 1) of the axis of the spectrometer. A disc was cut from the crystal of Fig. 8, with a face in the cross-section of the specimen, and was mounted as described above. Fig. 8 then corresponds

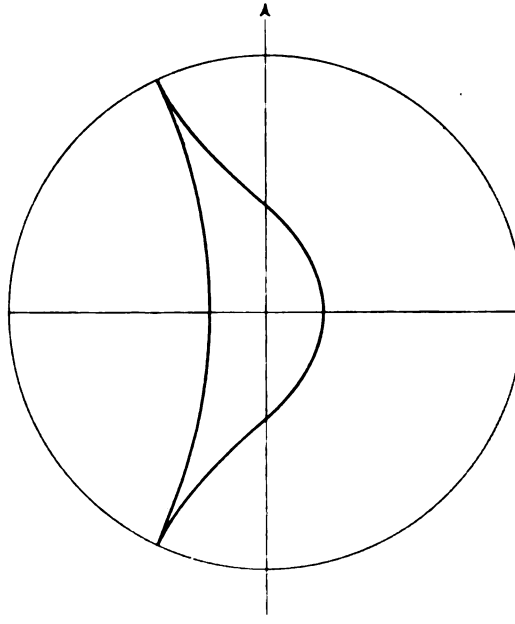


Fig. 10.

to a projection with a "horizontal" axis and may be used in conjunction with Fig. 10. In this way Fig. 11 has been drawn. Each of the circular bands of this diagram corresponds to one of the axes. When a reflection can be obtained for an axis with the specimen mounted so that a given line in the surface is pointing in the "upward" direction of the axis of rotation, the corresponding band has been blackened in the direction of that line. It again appears that the more steeply inclined axes are accessible only over very short ranges, whereas the axis "*G*" is now accessible for any mounting of the specimen. But every axis now falls at some

Digitized by Google

time within the region of accessibility. From such a diagram the most suitable orientation for mounting the specimen is readily selected.

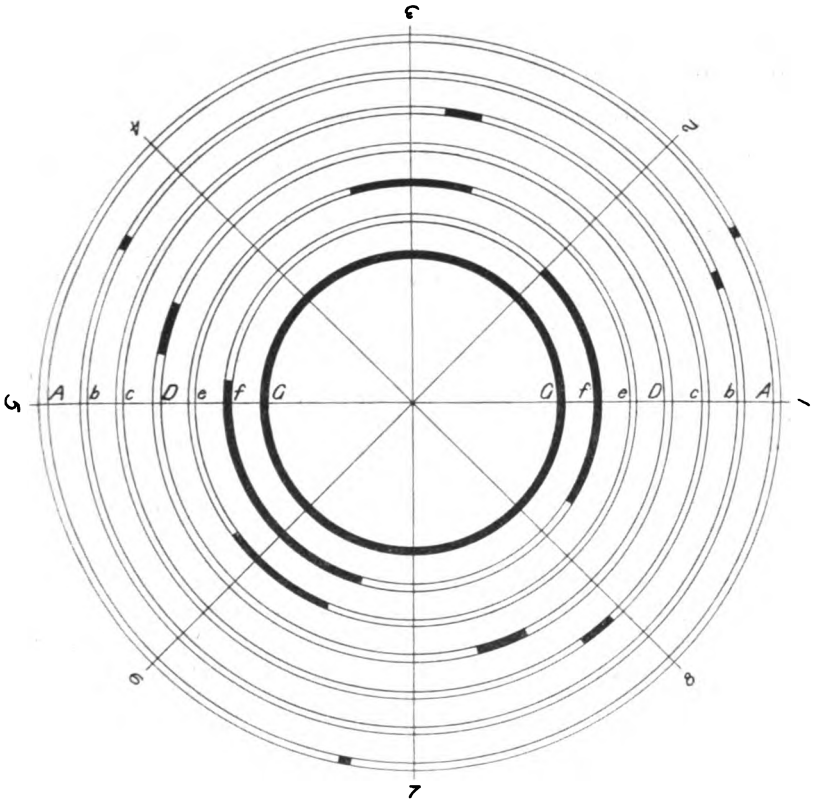


Fig. 11.

The preparation and use of such diagrams as are described above often save a great amount of labour and of fruitless work in problems involving the analysis of crystal orientations.

In conclusion, the author wishes to thank Professor Sir Ernest Rutherford and Professor G. I. Taylor for their interest in the work.

Digitized by Google

Contribution to the Theory of the Diffusion Pump. By L. WERTENSTEIN, D.Sc., Professor of Radiology at the Free-University, Warsaw. (Communicated by Professor Sir E. RUTHERFORD.)

[*Received* 12 November, *read* 22 November 1926.]

There are two effects involved in the sucking action of the diffusion pump. One is the penetration of gas molecules into the mercury vapour jet, and the second is the existence of a gradient of the partial pressure of the gas in the direction of the jet, so that the pressure at the origin of the jet is much lower than at the end of it.

The importance of the first effect depends on the size of what may be called the "slit" of the pump. It is really a slit in the original diffusion pump of Gaede and is, in modern pumps, a rather wide annular space between the base of the jet and an outer tube. In his fundamental paper* Gaede investigates the influence of the dimensions of the slit on the "speed of evacuation," *i.e.* the quantity of gas pumped in unit time by the jet. He calculates the rate of evacuation as a difference between the "gas coming into the jet" and the "gas coming back out of the jet," but he does not take into account the fact that this last effect is essentially influenced by distribution of velocities in the jet, which, instead of being isotropic, as he assumes, approaches, especially for powerful pumps, to a more unidirectional character.

The "sucking effect" is only possible when the partial pressure at the issuing jet is lower than the pressure of the gas in the vessel to be evacuated and the limiting pressure of the pump is due to the second effect to which we alluded above, *i.e.* the establishment of a pressure gradient along the jet. Let us investigate this point more closely. We consider a stream of mercury vapour in a cylindrical tube of length l. Let the concentration of Hg molecules at a distance x be n, their average velocity u. We neglect the variation of u along the radius of the tube section and we assume nu as constant through the length of the tube. These are different from actual conditions. The simplification we are making, while allowing for an exact calculation of the effect, does not alter the essential features of the phenomenon.

Suppose, now, the tube to be filled with a gas; there would be obviously a pressure gradient of this gas established in the direction of the jet. Gaede† and, apparently independently of him,

* *Ann. d. Phys.*, Vol. 46, p. 357, 1915.
† Gaede, *l.c.*

Digitized by Google

Hertz*, tried to represent this gradient in the following way: if dp_1/dx (we shall denote quantities relative to the gas by the index 1) is the gradient, there would be $\rho_1 D_1 (dp_1/dx)$ grammes of the gas streaming back per unit time and area, ρ_1 being the density of gas at unit pressure, and D_1 its diffusion coefficient at the point x. If conditions are steady, this quantity must be balanced by the quantity $\rho_1 p_1 u$, representing the mass of gas carried by the vapour stream per unit time and area. We get in this way

$$\frac{dp_1}{dx} = \frac{u}{D_1} p_1.$$

If we assume now that $D_1 n = D_0$ and is constant, which is approximately correct if there is a large excess of the mercury vapour, on account of condition $nu = \nu =$ constant, we obtain

$$\frac{dp_1}{dx} = \frac{\nu}{D_0} p_1,$$

which leads to an exponential distribution of pressure:

$$p_1 = p_0 e^{\nu x / D_0} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1).$$

The way of introducing the diffusion coefficient in this formula corresponds exactly to the dynamical method of calculating this coefficient used first by Langevin† and developed by Chapman‡ and by Enskog§. In this method one of the diffusing gases is supposed moving relatively to the other with a velocity u which may be put parallel to the x-axis. There is a transfer of momentum in the direction of the x-axis from the moving to the stationary gas. If the momentum transferred per unit time and volume is F, we get

$$F = dp/dx.$$

On the other hand,

$$dp/dx = \frac{pu}{D},$$

so that we have

$$D = \frac{pu}{F}.$$

The calculation leads, if the molecules are assumed to be elastic spheres, to the following value of the diffusion coefficient of the gas m into gas m_1.

$$D = \frac{3}{16 n_1 \sigma^2} \times \left(\frac{m + m_1}{m_1}\right)^{\frac{1}{2}} \left(\frac{2RT}{\pi M}\right)^{\frac{1}{2}}.$$

* Hertz, *Zeit. f. Phys.*, Vol. 14, p. 35, 1923.
† Langevin, *Ann. de Physique*, p. 275, 1905.
‡ Chapman, *Phil. Trans.* A, Vol. 211 p. 433, 1911.
§ Enskog, *Physik. Zeitschr.*, Vol, 12, p. 58, 1911.

Digitized by Google

This value is different from values obtained by using the notion of the free path. Now the experimental conditions described above offer a possibility of measuring directly dp/dx and of so giving a check to the dynamical theory.

We shall use for this comparison the experimental data obtained by Gaede in the following way.

A stream of mercury vapour of constant temperature was passed through a cylindrical tube connecting two vessels. The whole apparatus contained initially a gas at a pressure p. Under the influence of the vapour stream, the pressure in the vessel into which the mercury vapour streams, rises and a steady state is obtained. Determining the ratio of pressures in both vessels, corrected for the "thermal pressure" effect due to a difference of temperature between the vessels, Gaede puts this ratio equal to

$$e^{\frac{\nu}{D_0}L},$$

where ν and D_0 are defined as in formula (1), and L is the length of the tube. ν is determined by measuring the weight of mercury distilling in unit time, and in this way D_0 is obtained. Gaede calculates the values of D_0/n_0 where n_0 is the number of molecules in 1 c.cm. at atmospheric pressure and at temperature of 140° (temperature of the tube). This value, according to his ideas, represents the diffusion coefficient of Hg at 140° through the gas investigated. For hydrogen Gaede obtained the value 0·66. For air, as an average between two rather discordant values, 0·21.

Now on, the dynamical theory, $\frac{D_0}{n_0}$ is equal to

$$D = \frac{3}{16 n_0 \sigma^2 \sqrt{\pi \frac{M}{2RT} \cdot \frac{m_1}{m + m_1}}}.$$

According to Dushman (*International Critical Tables*, p. 92), the free path of a Hg atom at 25° and normal pressure is $3{\cdot}2 \cdot 10^{-5}$ cm., which gives for the diameter of the Hg atom $5{\cdot}67 \cdot 10^{-8}$ cm. For the hydrogen molecule the diameter is $2{\cdot}38 \cdot 10^{-8}$ cm. and for the average air molecule we may assume the value $3{\cdot}25 \cdot 10^{-8}$ cm. This gives:

$$\sigma_{\text{air-Hg}} = 4{\cdot}46 \cdot 10^{-8} \text{ cm.}; \quad \sigma^2 = 19{\cdot}5 \cdot 10^{-16} \text{ cm.}^2;$$

$$\sigma_{\text{H}_2\text{-Hg}} = 4{\cdot}025 \cdot 10^{-8} \text{ cm.}; \quad \sigma^2 = 16{\cdot}2 \cdot 10^{-16} \text{ cm.}^2.$$

Digitized by Google

At 140° n_0 is equal to

$$\frac{2{\cdot}7\,.\,10^{19}\,.\,273}{413} = 1{\cdot}785\,.\,10^{19},$$

so that we get, for H_2:

$$\frac{D_0}{n_0} = \frac{3}{16\,.\,1{\cdot}785\,.\,10^{19}\,.\,16{\cdot}2\,.\,10^{-16}\sqrt{3{\cdot}14\,\frac{200{\cdot}2\,.\,2{\cdot}016}{16{\cdot}6\,.\,10^{7}\,.\,413\,.\,202{\cdot}22}}}$$

$$= \frac{6{\cdot}5\,.\,10^{-6}}{9{\cdot}6\,.\,10^{-6}} = 0{\cdot}675 \quad \text{(Gaede 0·66 obs.)},$$

and for the air:

$$\frac{D_0}{n_0} = \frac{3}{16\,.\,1{\cdot}785\,.\,10^{19}\,.\,19{\cdot}5\,.\,10^{-16}\sqrt{3{\cdot}14\,\frac{200{\cdot}2\,.\,28{\cdot}8}{16{\cdot}6\,.\,10^{7}\,.\,413\,.\,229}}}$$

$$= \frac{5{\cdot}37\,.\,10^{-6}}{3{\cdot}4\,.\,10^{-5}} = 0{\cdot}158 \quad \text{(Gaede 0·205 obs.)}.$$

The agreement is fairly good for hydrogen and rather poor for air; but the two series of measurements from which we took the average, give respectively 0·17 and 0·24, so that our comparison is necessarily a rough one.

We may now try to apply this theory to the diffusion pump itself. We consider the part of this pump between the orifice of the mercury jet and the tube leading to the preparatory pump, *i.e.* practically the whole water-cooled part of the diffusion pump. It is easy to show that the ratio of "back" and "final" pressure calculated by means of the exponential formula (1) is much greater than 10^6, even if we take a very moderate estimate of the quantity of mercury distilling per unit time. It is, however, obviously useless to enter into details of these calculations, because the essential conditions involved in our theory, *i.e.* constancy of temperature and of number of Hg atoms crossing each section of the tube are very far from being fulfilled in the case of the diffusion pump. We may, however, verify in a simple manner the consequences of the theory. If the ratio of back and final pressure depends on the value of D_0, then it must increase with the molecular weight and the size of molecules of the pumped gas. In other words, the pumping effect must be higher, *e.g.* for CO_2 than for A and for A than for H_2. This consequence of the theory seems at first sight to be in contradiction with experiment, because it is generally assumed that the diffusion pump pumps out the lighter gases more quickly than the heavier. We must not forget that we are speaking here of the final ratio of pressures and not of the speed of evacuation which depends obviously on the law of flow of

Digitized by Google

gases and for this reason increases with decreasing molecular weight. It is not easy to measure the final ratio of "back" and "low" pressure because, as a rule, the low pressure is too small for any precise measurements. For this reason, in order to verify our theory, we have worked under such conditions that the final pressure produced by the pump was high enough to allow exact measurements. The best way of doing so is to run the pump at a comparatively low temperature. In our experiments a thermocouple was fixed in contact with the outside wall of the distillation tube of a Langmuir pump, as near to the orifice as possible. The thermocouple and, of course, the whole tube, were wrapped with

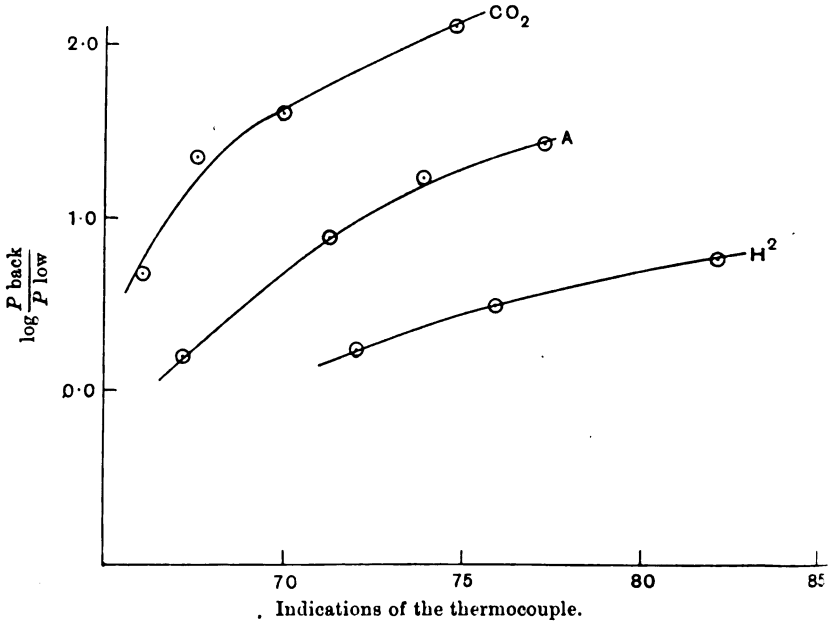


Fig. 1.

asbestos. It is clear that the thermocouple does not indicate exactly the temperature of the Hg stream, but that did not matter for our purpose, which was to *compare* the steady state obtained with different gases for a given speed of distillation in the diffusion pump. We may safely suppose that this speed must be always nearly the same when the indication of the thermocouple is the same, and when the pressures of the investigated gases are low and of the same order of magnitude in every experiment. The experiments have been carried out in the following way. The vessel to be evacuated consisted of a MacLeod gauge and some additional tubes, the volume of which was 300 c.cm. The "back pressure

Digitized by Google

vessel" was a bulb of about 6 litres capacity. The whole apparatus was first evacuated by means of a Gaede rotary pump and then separated from this pump. The gas under investigation was then introduced up to a pressure of a few hundredths of a millimetre of Hg with the diffusion pump cooled, so that the pressure indicated by the MacLeod gauge was the same in the whole apparatus. This initial pressure did not vary much in the back side, with the diffusion pump set in action, owing to the large volume of the back pressure bulb, so that on reading the MacLeod gauge we may easily deduce at every moment the ratio of "back" and "low" pressure. An electric current of given intensity was sent through the heating coils of the diffusion pump, and when the corresponding indication of the thermocouple became constant, two successive readings of the MacLeod gauge were taken in order to be sure that the state was a steady one.

The results are given in Fig. 1. In this figure the abscissae are the indications of the thermocouple and the ordinates the logarithms of the ratio of back and low pressure for H_2, A and CO_2 respectively. The initial pressures for these three gases were respectively $1{\cdot}85 \cdot 10^{-2}$, $1{\cdot}15 \cdot 10^{-2}$ and $1{\cdot}42 \cdot 10^{-2}$ mm. Hg. As is seen from the figure, the effect of the diffusion pump is, at the same temperature, higher for CO_2 than for A, and higher for A than for H_2. These three gases range themselves, in accordance with our theory, in the order of decreasing molecular weight and also of decreasing size of the molecule. For reasons discussed above, it seems useless to insist upon the comparison of numerical results with the theoretical expectation. The difference of behaviour of different gases is a very striking one and suggests the possibility of using the diffusion pump for separation of gases.

Digitized by Google

On the Addition of the Primary Aberrations. By Mr G. C. STEWARD, Gonville and Caius College.

[*Received* 15 November, *read* 22 November 1926.]

§ 1. In the application of the Characteristic Function of Hamilton, or of any allied function, to the computation of a symmetrical optical system three steps are necessary. The performance of the system as a whole must be considered, and from this it appears that the aberrations may be derived from certain 'aberration coefficients' which occur in the expansion of an 'aberration function'.* In the second place, relations must be obtained giving the properties of the complete system in terms of the properties of the component systems, which, in general, will be single refracting surfaces†; and finally, an evaluation of the coefficients must be made for the simple system—a single surface‡. For the first and third of these steps reference may be made elsewhere, and also for a general investigation of the second; a simple derivation is given, in the present note, of the necessary relations between a composite system and its components in the case where first order aberrations only are considered.

§ 2. We use a modification of the Characteristic Function, viz. the Eikonal, associated with the name of Bruns§. Let $u - \Phi$ be the modified and reduced eikonal, the base points being conjugate and given by a reduced magnification m; here u is the form of the eikonal giving perfect definition over the conjugate normal planes m, and Φ denotes the presence of aberrations for these two planes, so that Φ is the 'aberration function.' The form of the function u is to some extent arbitrary, but it is commonly made definite by imposing the condition that the axial points of the pupil-planes of the system shall be free from aberration‖. If then L, M, N and L', M', N' be the direction cosines of an incident and the corresponding emergent ray, and Y_1, Z_1 the co-ordinates of the ideal image of a point Y, Z, the axis of the system being in each case the axis of x, the aberration displacements D_{Y_1} and D_{Z_1}, produced by the system, are given by the relations

$$D_{Y_1} = \left(\frac{\partial}{\partial M'} + m\frac{\partial}{\partial M}\right)\Phi \quad \text{and} \quad D_{Z_1} = \left(\frac{\partial}{\partial N'} + m\frac{\partial}{\partial N}\right)\Phi \ldots (1);$$

* *Phil. Trans.* A. (225), 1925, Part I; *Trans. Camb. Phil. Soc.*, No. IX, 1926; *Proc. Opt. Convention*, 1926, Part II, p. 776.

† T. Smith, *Trans. Opt. Soc.*, Vol. XXV, No. 4, 1923-24.

‡ *Proc. Camb. Phil. Soc.*, Oct. 1926.

§ "Das Eikonal," *Abhandlungen der math.-phy. Classe der k. sächsischen Akad. d. Wissenschaften*, Bd. XXI., Leipzig, 1895; see also *Phil. Trans.* A. (225), *loc. cit.*

‖ T. Smith, *Trans. Opt. Soc.*, Vol. XXIII, No. 5, 1921-22.

Digitized by Google

exemplifying the fact that Φ gives all the aberrations of the system and depends upon these aberrations alone.

Now the function Φ, and indeed the eikonal in general for a symmetrical system, may be expressed in terms of three variables α, β and γ given by

$$\alpha d^2 = (M - sM')^2 + (N - sN')^2,$$

$$\beta d^2 = (M - mM')(M - sM') + (N - mN')(N - sN'),$$

and $$\gamma d^2 = (M - mM')^2 + (N - mN')^2,$$

d being the distance between the exit-pupil and the conjugate plane in the image space and s the magnification associated with the pupil planes; and if we consider only the first order aberrations we may write

$$8\Phi = \sigma_1\alpha^2 - 4\sigma_2\alpha\beta + 2\sigma_3\alpha\gamma + 4\sigma_4\beta^2 - 4\sigma_5\beta\gamma + \sigma_6\gamma^2,$$

the numerical coefficients being chosen for subsequent simplicity. Applying now (1), the aberration displacements from ideal imagery are given by

$$2d^3 D_{Y_1} = \sigma_1 S + \sigma_2 C + (\sigma_3 + 2\sigma_4) A + \sigma_5 D$$

and $$2d^3 D_{Z_1} = \sigma_1 S' + \sigma_2 C' + \sigma_3 A', \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2),$$

where $$\begin{matrix} S \\ S' \end{matrix} \equiv -\rho^3 \begin{matrix} \cos\phi \\ \sin\phi \end{matrix}, \quad \begin{matrix} C \\ C' \end{matrix} \equiv \rho^2 Y_1 \begin{matrix} (2 + \cos 2\phi) \\ \sin 2\phi \end{matrix},$$

$$\begin{matrix} A \\ A' \end{matrix} \equiv -\rho Y_1^2 \begin{matrix} \cos\phi \\ \sin\phi \end{matrix} \quad \text{and} \quad D \equiv Y_1^3.$$

Here ρ, ϕ are the polar co-ordinates of a point upon the exit-pupil—the point through which the emergent ray passes; further, we have written $Z_1 = 0$, which is always possible, without loss in generality, by a proper choice of the axis of z—so far arbitrary. It is evident, then, that the first order aberrations, i.e. spherical aberration, coma, astigmatism and curvature of the field and distortion are given completely by the *five* aberration coefficients $\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5$; but, in fact, we ought to include the sixth, σ_6, which governs spherical aberration for the pupil planes, to the first order.

§ 3. Consider now the addition of n symmetrical systems having the same axis—the suffix λ attached to any of the above quantities denoting the value of that quantity for system λ; let the combined magnifications for systems p to q inclusive be denoted by $m_{p,q}$ so that, for example,

$$m_{p,q} = m_p m_{p+1} \ldots m_q.$$

Suppose that a ray cut the $n+1$ conjugate planes, after passage through the various parts of the system, in the points Y_0, Z_0; Y_1, Z_1; ... Y_n, Z_n; then we have the following relations:

Digitized by Google

$$Y_1 - m_1 Y_0 = (D_{Y_1})_1$$
$$Y_2 - m_2 Y_1 = (D_{Y_2})_2$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$Y_n - m_n Y_{n-1} = (D_{Y_n})_n, \quad \cdots\cdots\cdots\cdots\cdots\cdots (1),$$

and the like equations for the Z co-ordinate. Further,

$$Y_n - m Y_0 = (D_{Y_1})_{1\ n} \quad \cdots\cdots\cdots\cdots\cdots\cdots (2),$$

so that, multiplying the relations (1) by the appropriate magnification and adding, we have

$$(D_{Y_n})_{1,n} = \sum_{\lambda=1}^{n} (D_{Y_\lambda})_\lambda m_{\lambda+1,n} \quad \cdots\cdots\cdots\cdots\cdots\cdots (3),$$

together with a corresponding relation involving Z in place of Y; and it is to be noticed that these relations are exact and not approximate. To the first approximation we have

$$Y_n = Y_\lambda m_{\lambda+1,n}$$

and also

$$\rho_n = \rho_\lambda s_{\lambda+1,n}$$

if $\ldots\rho_\lambda\ldots$ be the radii of the successive images of the aperture stop; whence substituting in (3), we have a relation involving various powers of Y_n and ρ_n which must be satisfied for *all* values of these quantities; so that we may equate to zero the algebraic sum of any one power of Y_n and of ρ_n. We may use (2) § 2, but it must be remembered that here we are dealing with modified distances, i.e. with reduced distances multiplied by the modified power $J(\equiv K/\mu\mu')$ of the system associated with them. Moreover, in the Y-equation we may write everywhere $\phi = 0$, and in the Z-equation $\phi = \pi/2$; then equating coefficients as indicated above and considering only the primary aberrations, we have

$$\left(\frac{\sigma_1}{J}\right)_{1,n} = \sum_{\lambda=1}^{n} \left(\frac{\sigma_1}{J}\right)_\lambda m^4{}_{\lambda+1,n}, \quad \left(\frac{\sigma_2}{J}\right)_{1,n} = \sum_{\lambda=1}^{n} \left(\frac{\sigma_2}{J}\right)_\lambda m^3{}_{\lambda+1,n} s_{\lambda+1,n},$$
$$\left(\frac{\sigma_3}{J}\right)_{1,n} = \sum_{\lambda=1}^{n} \left(\frac{\sigma_3}{J}\right)_\lambda m^2{}_{\lambda+1,n}\, s^2{}_{\lambda+1,n},$$
$$\left(\frac{\sigma_4}{J}\right)_{1,n} = \sum_{\lambda=1}^{n} \left(\frac{\sigma_4}{J}\right)_\lambda m^2{}_{\lambda+1,n}\, s^2{}_{\lambda+1,n},$$
$$\left(\frac{\sigma_5}{J}\right)_{1,n} = \sum_{\lambda}^{n}{}_{1} \left(\frac{\sigma_5}{J}\right)_\lambda m_{\lambda+1,n}\, s^3{}_{\lambda+1,n}, \quad \left(\frac{\sigma_6}{J}\right)_{1,n} = \sum_{\lambda=1}^{n} \left(\frac{\sigma_6}{J}\right)_\lambda s^4{}_{\lambda+1,n}$$
$$\cdots\cdots\cdots\cdots\cdots\cdots (4).$$

The first five of these relations follow by direct comparison of coefficients, the last by symmetry; for just as σ_1 governs spherical aberration for the axial points associated with reduced magnification m, so σ_6 governs axial aberration for the conjugate points

Digitized by Google

associated with reduced magnification s. We have substituted first approximations in (3); the addition relations for higher order aberration coefficients could be obtained by further approximation: but it is better, perhaps, to obtain them by a more general method*.

It is not convenient to calculate directly the quantities $\sigma_1 \ldots \sigma_6$ given above, for these all depend upon the conjugate planes m which are chosen and also upon the pupil planes s of the system, but simple relations may be obtained by means of (4), giving them in terms of other quantities which are invariants of the optical system, i.e. independent of m and s; and this may be effected by reference to the focal reduced eikonal of the system†.

* T. Smith, "The Addition of Aberrations," *Trans. Opt. Soc.*, Vol. xxv, No. 4, 1923–24.

† *Proc. Camb. Phil. Soc.*, Oct. 1926.

Digitized by Google

The absorption spectra of "saturated" and "unsaturated" organic substances. By Mr J. E. PURVIS.

[*Received* 15 October, *read* 8 November 1926.]

This investigation is to discuss some of the conditions which influence the absorption spectra of various "saturated" and "unsaturated" organic compounds. A condensed cadmium spark was used as the source of radiant energy; alcoholic solutions of the substances were employed, except for stearic acid, which was dissolved in ether. The vapours were examined in a quartz tube 200 mm. long, placed in a bath which could be heated to various temperatures.

Cyclo-hexane. Hartley and Dobbie (*Journ. Chem. Soc.*, 1900, vol. 77, p. 846) showed that solutions of this substance had no selective absorption when the radiant energy was passed through various thicknesses of different strengths. Stark and his co-workers (*Jahrb. Radioak. u. Electr.*, 1913, vol. 10, p. 139), using a fluorspar spectrograph, found that the rays were transmitted to $\lambda 1850$ without any absorption, when they were passed through the *vapour* of the substance in a tube 200 mm. long, and under a pressure of 56 mm. The author has examined alcoholic solutions, through various thicknesses, of M, $M/10$, and $M/100$ strengths, but no bands appeared. In a $M/100$ strength, and through 28 mm. thickness, the rays were transmitted to $\lambda 2310$, and through 2 mm. thickness to $\lambda 2190$.

Cyclo-hexene. Hartley and Dobbie (*loc. cit.*) state that solutions of this substance showed greater general absorption than cyclohexane. Stark and Levy (*loc. cit.*, p. 179) in the *vapour* found, under the reduced pressures of 30 mm. and of 1 mm., two bands at about $\lambda 2200$ and $\lambda 1900$. The author has examined solutions of M, $M/10$, and $M/100$ strengths, when the rays passed through varying thicknesses of from 2 mm. to 28 mm., but no bands were observed. For example, in the $M/100$ solution, the rays were transmitted through 2 mm. thickness to about $\lambda 2130$; through 10 mm. thickness to $\lambda 2150$; through 20 mm. to $\lambda 2190$, and through 28 mm. to $\lambda 2240$. The vapour of the substance was also examined in a 200 mm. tube, at temperatures varying from 9° to 50°, but no band at about $\lambda 2200$ was observed; the rays were transmitted at 9° to about $\lambda 2150$, and at 30° to about $\lambda 2190$. The non-appearance of the vapour bands, observed by Stark and Levy at the considerably reduced pressures, may be a result of the influence of the solvent in the case of the solutions, and by the much higher pressures in the case

Digitized by Google

of the vapours, phenomena which have been observed in other vapours and their solutions. Stark and Levy argue that the vapour bands of hexene have their origin in the unsaturated condition, or the "ethylenic" linking, of the molecule.

Cyclo-hexanol; o-, m-, and p-methyl cyclo-hexanol. Solutions of these four substances of different strengths ranging from *M*, *M*/10, and *M*/100, and through thicknesses ranging from 2 mm. to 28 mm., were examined but no bands were observed. The solutions were very transparent; for example, the rays were transmitted to λ2190 through 2 mm. of a normal solution; to λ2270 through 10 mm.; to λ2370 through 20 mm.; to λ2530 through 28 mm. The vapours were also examined in a 200 mm. tube at temperatures from 50° to 100°; and, like the solutions, they were very transparent.

Cyclo-hexyl formate. Different strengths of solutions of this substance were also examined, but no bands were observed. The solutions were very transparent. For example, in a *M*/100 solution the rays were transmitted to λ2120 through 2 mm. thickness; to λ2150 through 10 mm.; to λ2240 through 20 mm.; and to λ2260 through 28 mm.

Cyclo-hexanone. Henderson, Henderson and Heilbrun (*Ber. Chem. Ges.*, 1914, vol. 47, p. 876), and Bielecki and Henri (*Ber.*, 1914, vol. 47, p. 1690), showed that this substance exhibited a large ketonic band at about 1/λ3600 (λ2777), and a similar band was found in the three *o-*, *m-*, and *p*-cyclo-hexanones. The author has repeated these observations and confirms them. He has also examined the *vapour* of cyclo-hexanone at temperatures ranging from 100° to 200° C., and finds a large band similar to that of the solution band. There is no division of this large band into a series of narrower ones.

Geraniol, geranyl acetate, geranyl butyrate, geranyl valerate, geranyl pelargonate. The author (*Journ. Chem. Soc.*, 1924, vol. 125, p. 406) examined solutions of geraniol of various strengths, but no bands were observed. Stark (*loc. cit.*) has examined the vapour of geraniolene, the unsaturated hydrocarbon from which geraniol may be obtained, and found, under a reduced pressure of 1–2 mm., a weak band in the neighbourhood of λ2300–λ2050, and a second band at λ1800; the rays did not extend beyond λ2400 when the vapour was examined under a pressure of more than 5 mm. Stark ascribes these bands to the "ethylenic" linking of the molecule.

By weighting a molecule with a constituent which itself gives no specific absorption, the absorption band (or bands) belonging to the original substance is often shifted towards the less refrangible regions. With this in view, the acetate, butyrate, valerate and pelargonate of geraniol were compared with geraniol itself. The acids of these esters gave no absorption bands, and the esters

Digitized by Google

themselves were also very transparent. Comparing the heavy pelargonate with geraniol, the following table gives the position beyond which the rays were absorbed:

		Geraniol	Geranyl pelargonate
M/10	28 mm.	λ2980	λ3230
	20 „	2970	3180
	10 „	2950	3080
M/100	28 „	2720	3040
	20 „	2650	3030
	10 „	2550	3000
M/1000	28 „	2250	2500
	20 „	2180	2490
	10 „	2140	2450

Solutions of oleïc, pelargonic, and stearic acids were also examined; oleïc acid has one unsaturated centre; pelargonic and stearic acids are both saturated acids. None of them showed any specific absorption. The following table compares the position where the rays were absorbed by various strengths and through different thicknesses of *M*/10 solutions:

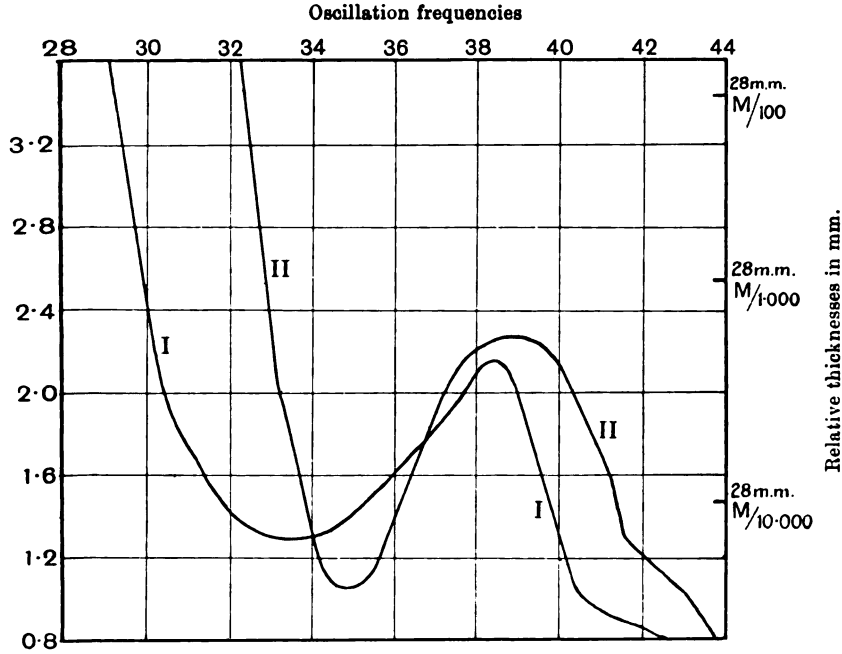
		Oleïc acid $CH_3(CH_2)_7 . CH : CH(CH_2)_7 . COOH$	Stearic acid (in ether) $CH_3(CH_2)_{16} . COOH$	Pelargonic acid $CH_3(CH_2)_7 . COOH$
M/10	28 mm.	λ3150	λ3000	λ2730
	20 „	3100	2850	2570
	10 „	3000	2500	2480

Pentabromo-phenol; tetrabrom-o-cresol. Solutions of these substances were examined, and both exhibited bands comparable with that of phenol, described by Hartley, Dobbie and Lander (*Journ. Chem. Soc.*, 1902, vol. 81, p. 929). The curves (Fig. 1) show that the band of the former compound is wider than that of the other, but the original phenol band has not been destroyed in replacing hydrogen atoms by heavier bromine atoms; there is a shift towards the red end caused by the weighting of the molecule, and there are differences in the intensities.

But specific absorption in the ultra-violet regions is not always governed by the "unsaturated" condition of the atoms of a

Digitized by Google

molecule. For instance, the author (*Journ. Chem. Soc.*, 1911, vol. 99, pp. 811, 1699, 2318) has shown that the vapours, and the alcoholic solutions, of iodo-derivatives of benzene and of toluene showed no specific absorption, whereas there were well-defined bands in the corresponding chlorine and bromine compounds. The entrance of a single iodine atom destroys specific absorption in the ultra-violet regions, although the iodobenzene remains "unsaturated" in the sense that it contains five chemically unsaturated valencies. In



I. Pentabromo-phenol
II. Tetrabrom-o-cresol

Fig. 1.

the infra-red regions, on the other hand, the author has described bands of iodobenzene between 2μ and 12μ which are fairly comparable with the phenomena of benzene and bromobenzene in those regions (*Proc. Camb. Phil. Soc.*, 1923, vol. 21, p. 556). Similarly, numerous narrow bands found in the vapours of *o*-, *m*-, and *p*-chlor- and bromanilines, were absent in the corresponding iodoanilines; and, like results appeared when comparing *p*-iodophenol with *p*- chloro- and *p*-bromophenols (*Journ. Chem. Soc.*, 1913, vol. 103, p. 1638).

Digitized by Google

It may also be recalled that Coblentz found bands in the infra-red regions between 2μ and 12μ both in saturated, as well as in unsaturated substances; for example, in hexane, ethane and ethyl alcohol: in acetic, valeric, and stearic acids and in other saturated compounds. (*Investigations of infra-red spectra*, Part I. 1905. Carnegie Institution of Washington Publications, No. 35.)

PUBLIC HEALTH CHEMICAL LABORATORY,
CAMBRIDGE.

Digitized by Google

Surface adsorption and the velocity of chemical action at gas-solid interfaces. By Dr F. HURN CONSTABLE, *Fellow of St John's College.*

[*Received* 16 November 1926, *read* 31 January 1927.]

It has been shown to be a necessary consequence of the Langmuir-Frenkel theory of the behaviour of gaseous molecules in collision with solid surfaces that chemically inert* diluents in the reactants should lower the rate of catalysis at the surface†. An expression for the fractional reduction in the reaction velocity has been obtained, and the purpose of this paper is to describe experimental tests of its validity.

An apparatus for the investigation of the effects of mixed vapours on catalytic activity has already been described‡. The effect of the addition of the vapours of water, benzene, and acetone on the initial stages of the dehydrogenation of alcohol by copper was investigated in the apparatus. The liquid reactant and diluent were contained in two glass boilers, and the amounts of the liquids vaporised were carefully controlled by the current supplied to the electric heating coils. In this manner the composition of the mixture could be varied within wide limits.

Experimental difficulties.

When water was used as the diluent an unexpected complication developed. The nichrome heating wires reacted with the water, giving hydrogen and leaving the hydroxides of the metals suspended in the water. When the rate of boiling was large this evolution of hydrogen was much more rapid than the dehydrogenation of the alcohol and constitutes a very serious source of error for which correction was made. A platinum heating coil was made and used in a boiler of special construction (Fig. 1) There was no evolution of hydrogen from this boiler even at the greatest rates of boiling, and the results confirmed those obtained with the nichrome heating coil when the latter were corrected for the hydrogen evolution from the water.

Palmer§ compared the rates of reaction of alcohol-water mixtures containing 4–30 per cent. of water with the pure alcohols on a reproducible catalyst, and reported that the water acted as a promoter, and further explained the phenomenon on the radiation

* I.e., that do not form complexes with the reactant on the surface.
† Constable, *Proc. Camb. Phil. Soc.*, vol. 23, p. 172 (1926).
‡ Constable, *Proc. Camb. Phil. Soc.*, vol. 22, p. 738 (1925).
§ Palmer, *Proc. Roy. Soc.* A, vol. 99, pp. 423–425 (1921).

Digitized by Google

theory. He used a nickel heating coil, but made no correction for the evolution of hydrogen due to the action of the aqueous alcohol on nickel under these conditions*. In the present work correction was always made for this source of error, and it was invariably

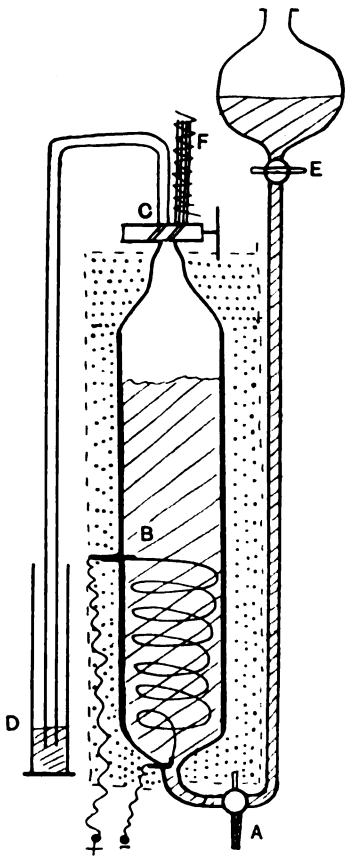


Fig. 1. The boiler.

The heating of the liquid is done by the bare platinum wire spiral sealed into the glass by means of much wider leads. This heating of bare wire is essential for rapid experimenting, and is very convenient. Immediately on increasing the heating current the vaporisation rate rises to the equilibrium value, and at the instant the current is cut off the vaporisation stops. Thus the effect produced can be followed at once, and any considerable lag in reaction velocity would be immediately appreciable.

The tap *E* enables the boiler to be evacuated for cleaning, and the liquid run in.

The tap *A* enables the boiler to be emptied after an experiment.

At *C* there is a two-way tap which communicates both with the measuring jar *D*, and the reaction vessel through the superheated capillary tube *F*. By distilling the liquid into *D*, the rate of vaporisation can be connected with the current passing, and any rate of vaporisation desired can be obtained.

found that water exerted a depressing action on the reaction velocity at the copper surface. With the other liquid diluents the depressing action was also found. Palmer's results were thus all vitiated by the spurious evolution of hydrogen from the heating coil.

* Mr Palmer agrees that this source of error was present in his experiments, for he observed the formation of considerable amounts of nickel hydroxide. Slight variations in the activity of the catalyst would be very serious in experiments of this type and constitute an additional source of error.

Digitized by Google

Investigation was first made to confirm that the rate of passage of the alcohol over the copper had little effect on the rate of reaction. The results are shown in Table I.

It is observed that the reaction velocity falls off slowly as the rate of passage of alcohol increases. Assuming 0·05 as the solubility of hydrogen in alcohol, we find that 7·3 c.c. of alcohol could

TABLE I. *The effect of rate of passage of vapour on reaction velocity.*

Temperature = 265° C.

Rate of distillation of alcohol c.c. per min.	Volume of hydrogen evolved in 5 mins.
0·61	22·0
	22·3
	22·2
	21·9
1·28	24·1
	22·3
	23·3
	22·3
2·00	21·1
	21·4
	21·4
	21·6
	21·9
3·41	19·2
	19·3
	19·1
	19·6
	19·0
3·74	19·5
	19·1
	19·7
	20·7
7·3	17·3
	17·5

absorb 0·36 c.c. or the maximum amount of hydrogen that could be absorbed by the alcohol distilling in 5 minutes is 1·8 c.c. If the results be corrected in this manner then there is still a slight fall in the reaction velocity which must therefore be due to a cooling effect which is exerted more on the catalyst surface than on the thermocouple. It is found to be somewhat difficult to keep the temperature of the catalyst constant with very high speeds of distillation. The evidence is therefore in favour of the true reaction velocity being independent of the rate of flow of the reacting

Digitized by Google

vapour when this exceeds some limit. When the speed of distillation does not exceed this limit the reaction velocity falls with the rate of passage of vapour. The effect is reversible, for on increasing the rate of flow the reaction velocity will again attain its terminal value. The explanation is that the reaction products are present in quantity comparable to the amount of alcohol present, and they are therefore adsorbed on the surface and reduce the surface of the catalyst that is active. In the table given above the partial pressure of the aldehyde and the hydrogen together never exceed 2 per cent of the total pressure, and the surface is therefore always covered completely with alcohol.

The effect of water as a diluent.

Rate of distillation of alcohol = 0·31 c.c. per minute.

Temperature = 240° C.

Rate of distillation of water c.c. per min.	Reaction velocity c.c. H_2 per min.	
	alcohol on	alcohol off
0·00	1·19	
	1·15	
	1·17	
0·24	0·61	
	0·61	
	0·60	
0·39	0·56	
	0·55	
	0·57	
0·61	0·64	
	0·64	
1·39	0·74	0·49
	0·73	
	0·74	
2·89	1·16	1·04
	1·17	
	1·16	
5·16	1·85	1·82
	1·88	

The solubility of hydrogen in water is only about 0·02, so that when a large excess of water is used as a diluent there will be little correction to be made for the dissolved hydrogen. The following method of experiment was therefore tried with satisfactory results. A constant rate of distillation was used, and the rate of boiling of the water continuously increased. After each increase in the amount of water used as diluent, the water stream

Digitized by Google

Rate of distillation of alcohol = 0·31 c.c. per minute.
Temperature = 250° C.

Rate of distillation of water c.c. per min.	Reaction velocity c.c. H_2 per min.	
	alcohol on	alcohol off
0·00	1·76	0·00
0·40	0·82	
1·00	0·77	
2·80	1·18	1·00

was completely cut off, and the activity of the catalyst measured to see if any irreversible change had taken place. The water stream was also completely cut off and correction made for the hydrogen evolved from the water boiler. The results were checked by using the boiler with the platinum heating coil. In all cases the rate of reaction with a mixture of gases was found to be constant at constant temperature, and there was no evidence that any appreciable time was taken to displace the film adsorbed previously.

TABLE II. *Summary of results with water.*

Fractional partial pressure of alcohol	Fractional reaction velocity	Fractional reaction velocity calculated from $\frac{1}{v} = \cdot 64 + \frac{\cdot 36}{p}$
1·00	1·00	1·00
0·56	0·83	0·78
0·41	0·58	0·66
0·26	0·45	0·49
0·20	0·41	0·41
0·197	0·38	0·39
0·087	0·26	0·22
0·038	0·11	0·10
0·032	0·10	0·09
0·024	0·06	0·07
0·019	0·05	0·05

In one experiment the catalyst was allowed to remain in a stream of water for 30 minutes. When the alcohol was turned on the reaction commenced immediately, and, as far as the measurements could tell, attained its maximum velocity at once. Such behaviour seems to be in strong contrast to that of water films *in*

Digitized by Google

vacuo which are extremely difficult to remove, and it is suggested that the difficulty in removing these films is due to the absence of molecular bombardment at high vacuum. When the diluent molecule can occupy the surface the chance of the slight pressure of water retaining the film complete is small since many of the diluent molecules are adsorbed.

The results obtained above were plotted on a curve (Fig. 2) and the results obtained with the boiler with the platinum heating coil are seen to be in good agreement with the others. From

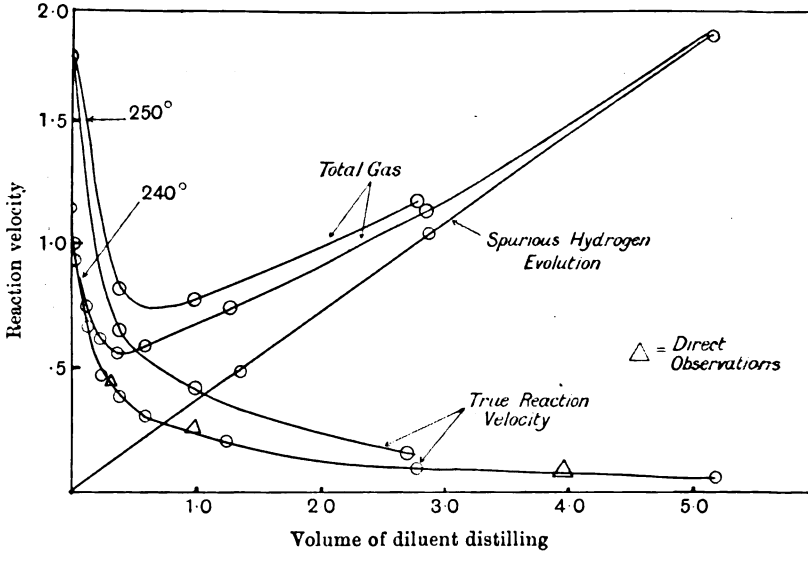


Fig. 2. Water as a diluent at constant temperature.

these curves the respective partial pressures and reaction velocities were read off, and the fractional partial pressures, and reaction velocities measured. These are compared with the calculated values in Table II. The agreement is good over a range of pressures of alcohol in water varying fifty times.

From the formula previously deduced

$$\lambda = \sqrt{\frac{M_a}{M_w}}\frac{\tau_w}{\tau_a}; \quad \therefore \frac{\tau_w}{\tau_a} = \sqrt{\frac{18}{46}} \times \cdot 36,$$

and hence the mean life of the water molecule relative to the alcohol molecule at 240° C. is 0·225.

Digitized by Google

The effect of acetone as diluent.

Purification of acetone. The bisulphite compound of acetone was decomposed with caustic soda, and distilled. The distillate was redistilled through a Dufton column, dried with calcium chloride, and finally redistilled through the Dufton column. The sample was sufficiently pure for the purpose since only a very slight poisoning effect was observed on the copper catalyst. Owing to the volatility of the acetone, extra resistances had to be added to reduce the current through the heating coil. There was no evolution of hydrogen from the boiler, and the experiments were in every way straightforward. The results are tabulated below.

TABLE III. *Acetone.*

Temperature = 240° C.

Fractional partial pressure of alcohol	Fractional reaction velocity observed	Fractional velocity calculated $v=p$
1·00	1·00	1·00
0·57	0·55	0·57
0·34	0·30	0·34

The reduction in the reaction velocity is linear with the pressure, and we have the case in which the molecular weights and the mean lives of the two molecules on the surface are the same.

The effect of benzene as diluent.

Purification of benzene. Pure commercial thiophene-free benzene was twice frozen. A constant freezing point of 5·5° C. was obtained which was independent of further crystallisation. It was found, however, that this benzene acted as a powerful poison, and the results were therefore rather laborious to obtain. The catalyst had to be oxidised and reduced after each experiment, and in addition only the experiments with low benzene concentrations gave any appreciable initial activity which could be extrapolated with moderate accuracy. The experiments represent the best attempt that could be made under the circumstances. It is interesting to note that Pease and Purdum* have observed this poisoning effect of benzene on copper. The benzene is converted into diphenyl, and simultaneous deactivation of the catalyst occurs.

* Pease and Purdum, *J. Amer. Chem. Soc.*, vol. 47, p. 1435 (1925).

Digitized by Google

TABLE IV. *Benzene results extrapolated.*

Temperature = 240° C.

Fractional partial pressure of alcohol	Fractional reaction velocity	Fractional reaction velocity calculated $\frac{1}{v} = \frac{4{\cdot}02}{p} - 3{\cdot}02$
1·00	1·00	1·00
0·90	0·65	0·69
0·81	0·50	0·52

The mean life of the benzene molecule is about five times that of the alcohol molecule.

The effect of diluents on the temperature coefficient of chemical change.

Let it be supposed that the catalyst is composed of a homogeneous distribution of centres of activity, and let $\Sigma\epsilon$ be the true heat of activation, that is the chemical work necessary to decompose the molecule in the adsorbed film, and μ_0' the heat of desorption of the unchanged molecule of reactant. Let the letters with double prime indicate the diluent molecules. Then the equation of reaction becomes

$$v = \frac{K \cdot e^{-\frac{\mu_0' + \Sigma\epsilon}{RT}}}{\left[1 + \lambda\left(\frac{1}{p} - 1\right)\right]\tau_0'}, \quad \lambda = \sqrt{\frac{\overline{M}_1}{\overline{M}_2}}\frac{\tau_0''}{\tau_0'} \cdot e^{\frac{\mu_0'' - \mu_0'}{RT}}, \quad \ldots\ldots(1)$$

and is also true for any exponential distribution of centres of activity provided that λ is a constant independent of the surface fields.

The quantity $\frac{1}{v}\frac{\partial v}{\partial T}$ is the true measure of the temperature coefficient of the reaction, and is nearly equal in magnitude to the ordinary 10° C. temperature coefficients used at the present time. The equation

$$v_T = Ae^{-Q/RT}$$

gives

$$\frac{1}{v}\frac{\partial v}{\partial T} = \frac{Q}{RT^2},$$

while

$$\frac{v_{T+10}}{v_T} = \frac{Q}{RT(T+10)}.$$

Digitized by Google

$$\text{From (1)}\quad \frac{1}{v}\frac{\partial v}{\partial T}=\frac{1}{RT^2}\left\{\frac{+\lambda(\mu_0''-\mu_0')\left[\frac{1}{p}-1\right]}{\left(1+\lambda\left[\frac{1}{p}-1\right]\right)}+\mu_0'+\Sigma\epsilon\right\}.$$

There are many important deductions to be made from this equation. When λ is small very little change is produced, and the more diluent present the less the effect. If the heat of desorption of the reactant is less than the diluent then the temperature coefficient is increased, if the two are equal then no effect is produced under any conditions, and if the heat of desorption of the reactant is greater than the diluent then the temperature coefficient is reduced.

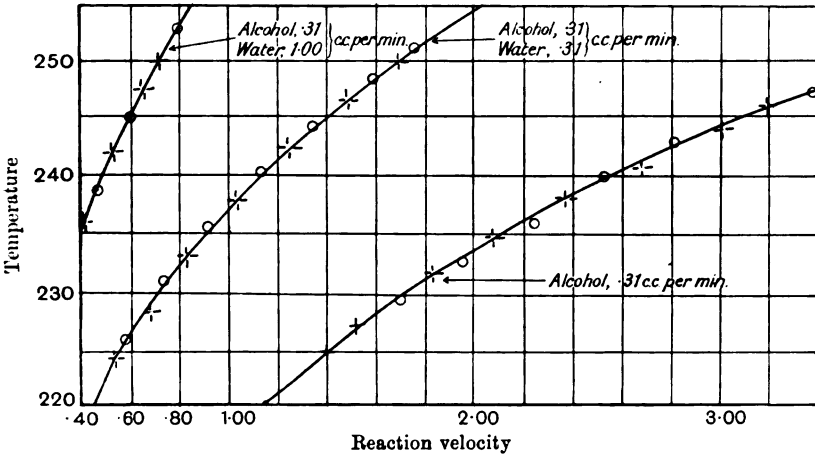


Fig. 3. The effect of water as diluent on the temperature coefficient.

When p is very small the equation may be written

$$\frac{1}{v}\frac{\partial v}{\partial T}=\frac{1}{RT^2}(+\mu_0''+\Sigma\epsilon), \quad \text{.....(2)}$$

and the alteration of temperature coefficient is independent of the amount of diluent present, though if the heats of desorption are not very different this effect cannot be traced experimentally. It is interesting to point out that the conditions for observing the effects of equation (2) are with low pressures of a powerful poison, and we see that this is a hitherto unsuspected cause of error in temperature coefficient measurement.

The effect of various concentrations of water on the temperature coefficient of dehydrogenation has been observed, but the alteration if present is slight as is shown by Table V.

Digitized by Google

TABLE V. *The reaction velocities for alcohol-water mixtures at various temperatures.*

Temperature	Reaction velocities in c.c. per min. and fractions of velocities for pure alcohol				
	Alcohol 0·31 c.c. per min.	Alcohol 0·31 Water 0·31 c.c. per min.	Fraction of initial activity	Alcohol 0·31 Water 1·00 c.c. per min.	Fraction of initial activity
225	1·40	0·56	0·40	—	—
230	1·75	0·73	0·42	—	—
235	2·15	0·92	0·43	0·41	0·19
240	2·59	1·12	0·43	0·50	0·19
245	3·12	1·37	0·44	0·61	0·20
250	3·95*	1·65	0·42	0·72	0·19
255	4·79*	2·01	0·42	0·86	0·21

* These values are extrapolated from the observations at lower temperatures.

It can be seen that the fraction of the surface covered by the alcohol depends far more on the partial pressure of the alcohol in the water vapour than on the temperature; in fact, the accuracy of the measurements is such that the fraction covered is constant within the limits of experimental error. This would be expected as a small value of $\lambda = \cdot 36$ has already been found, and the chemical similarity of water and alcohol together with the short range of molecular attraction and similar orientation of the adsorbed molecules, would lead us to expect heats of desorption which are not greatly different. The alteration in temperature coefficient is therefore small.

Results with thoria as a catalyst for dehydrogenation.

With thoria both dehydration and dehydrogenation of the alcohol occur, so that it is necessary to apply a substantial correction for the ethylene contained in the evolved gas. Both reactions were separated and it was found that, while the dehydrogenation followed the law proposed previously, the dehydration showed a marked divergence from it, and the curve inflected for small concentrations of water as diluent. It is suggested that a complex with water is formed and that this complex is much more reactive than the adsorbed alcohol alone. Thus, a curve showing a maximum is superposed upon the law previously discussed; and

Digitized by Google

in special cases the diluent may act as a promoter in small concentrations at the start, becoming a poison as the concentration increases. There seem to be insufficient data to work this case out completely at present. The following figures show that the dehydrogenation follows the reciprocal law with some accuracy.

TABLE VI.

Fractional pressure of alcohol in alcohol water mixture	Calculated velocity $\frac{1}{v} = \frac{5{\cdot}3}{p} - 4{\cdot}3$	Observed velocity
1·00	1·00	1·00
0·80	0·431	0·44
0·60	0·221	0·20
0·40	0·115	0·12
0·20	0 045	0·05

Results obtained by the use of the coherer.

Palmer* has developed an approximate method of estimating the composition of the adsorbed film by observing the change in the potential difference required to desorb a gas film from a mixture of gases when it is allowed to stand. He explains that the initial film has exactly the same composition as the gas, but that as the time proceeds the mean lives of some of the molecules are so much longer than that of the others that the film becomes much richer in these and the composition changes to a limiting value. He has performed his experiments at room temperature, and the time taken for the equilibrium to be reached can be conveniently measured in hours, which time is comparable with the mean life of a molecule under these conditions.

Such behaviour is in very marked contrast to that occurring at 250° C, when, so far as my observations could say, the reduction in reaction velocity when water is admitted to an alcohol stream passing over copper is immediate.

The results for carbon monoxide-oxygen mixtures correspond with the formula if the errors of the method are also taken into consideration, for, owing to the shapes of the voltage-composition curves, very considerable errors occur round the values 0·2 and 0·4 for the composition of the gaseous phase. At these values both voltage curves† pass through a maximum or a minimum, and it

* Palmer, *Proc. Roy. Soc.* A, vol. 106, p. 55 (1924); vol. 110, p. 133 (1925).
† *Loc. cit.* p. 137.

Digitized by Google

would seem that altering the gas composition from 0·4 to 0·5 has no effect on the composition of the final film. This point is not discussed by the author, and the value 0·4 is omitted from his table. As can be seen below, the agreement is good except in this region, and it seems reasonable to assume that the error is experimental and inherent in the method.

TABLE VII. *Composition of the film in molecular fractions of carbon monoxide.*

Initial	Equilibrium	Calculated $\frac{1}{n} = 0{\cdot}614 + \frac{0{\cdot}386}{p}$
0·05	0·125	0·120
0·10	0·20	0·224
0·20	0·30	0·394
0·30	0·46	0·526
0·40	—	0·632
0·50	0·70	0·728
0·60	0·80	0·805
0·70	0·86	0·866
0·80	0·92	0·923
0·90	0·97	0·964
1·00	1·00	1·000

Discussion of results.

When the centres of activity of a surface may be regarded as homogeneous, and the reacting substance forms no complexes with the diluent when in the adsorbed layer, then the formula *

$$\frac{1}{v} = (1-\lambda) + \frac{\lambda}{p}, \quad \lambda = \frac{\tau_0'' \cdot e^{\mu_0''/RT}}{\tau_0' \cdot e^{\mu_0'/RT}} \left(\frac{M_1}{M_2}\right)^{\frac{1}{2}}$$

is found to agree closely with the facts for a number of surface films.

In the deduction of this equation it is essential to assume that reaction takes place while the reactant molecule is in the adsorbed state. Langmuir and Frenkel suggested that each molecule that hit the surface was condensed and remained during the time interval τ on the surface, and then evaporated. This mechanism gives a good interpretation of the facts as regards the evaporation and condensation of the films, but gives no clue to the state of the molecules while they are adsorbed.

* Constable, *Proc. Camb. Phil. Soc.*, vol. 23, p. 176 (1926).

Digitized by Google

The film formed might be a two-dimensional gas, liquid, or solid. In the first two states the adsorbed molecule would be free to wander over the surface in an erratic manner before it evaporated, while in the latter case it would remain during its mean life upon the spot on which it first settled.

There are two possible methods by means of which chemical reaction could occur between different molecules. It might be that the collision of fast moving molecules capable of motion in the film produces chemical change, in which case the reaction velocity would be proportional to the product of the concentrations of each species in this film, which is proportional to the number of collisions; or that each species has to be absorbed side by side simultaneously on particular spots on the catalyst, namely the centres of activity of the surface. Both mechanisms give rise to the same form of equation, hence we may work out the dynamics of surface action by either means and obtain agreement with experiment, and not have advanced our knowledge of the mechanism of reaction in the film at all.

The observations of Pease and Stewart, showing that very small concentrations of poison will reduce the reaction velocity to a negligible value, are not so decisive in favour of active centres as they appear at first. It is possible that poison acts as a negative catalyst in the film, in which case measurements of this type would be valueless in measuring the fraction the active surface bears to the total.

The most conclusive argument against surface catalysis being caused by molecular bombardment of the molecules against each other in the surface film lies in the fact that adsorption is very general for many substances, but that catalysis is rare, and very specific to one substance, depending both on the chemical nature and the physical state of the catalyst. Apart from the arrangement of the surface atoms leaving external fields of sufficient magnitude to ensure condensation of the bombarding gas molecules, there is also another property which they possess of making chemical action possible. A preliminary activation is necessary before molecules can react, and, owing to the short range of molecular fields of force, it seems essential that reaction should occur in the molecules actually residing on the centres of activity of the surface. They are momentarily part of the lattice structure, and it therefore seems that it is in the molecules that remain adsorbed for most of their mean life on the centres of activity that activation occurs.

It seems, therefore, that we have an explanation of the mystery that adsorption is general, catalysis rarer. If the molecules form a gaseous or liquid film and the mobility of the molecules is considerable, the probability of activation is small, whereas, if the

Digitized by Google

adsorbed molecule becomes momentarily part of a solid structure, activation is possible when the electrostatic and electromagnetic fields of force forming the centres of activity of the service are suitable, and in this case the adsorbing surface becomes an active catalyst.

Summary.

The relation between the reaction velocity of catalytic dehydrogenation and the rate of passage of alcohol vapour over the catalyst has been found by experiment. At low rates of flow the concentration of the aldehyde and hydrogen in the effluent gas, and therefore over the catalyst, is comparable with that of the reactant. The whole of the surface is thus not covered with the reactant, and the adsorbed reaction products cause marked decrease in the reaction velocity.

When the concentration of the reaction products becomes sufficiently small, that is, when the rate of passage of the alcohol vapour exceeds some limit, the velocity is approximately independent of the rate of flow of reactant. With very high speeds the reaction velocity tends to fall again, but this is probably due to the difficulty of keeping the temperature of the catalyst constant.

The formula previously deduced connecting the partial pressure of the reactant in a mixture with the fractional reaction velocity has been verified for mixtures of alcohol vapour with water, acetone and benzene; and the composition of the adsorbed gas film with the fractional partial pressure of the gases for mixtures of carbon monoxide with oxygen on platinum.

The effect of mixture of vapours on the temperature coefficient of reaction is considered, and equations deduced to show the alteration produced. If the heat of desorption of the reactant is greater than that of the diluent, then the temperature coefficient is diminished, if they are equal there is no alteration, and if less the temperature coefficient is increased. Experiments show that water present in alcohol has little effect on the temperature coefficient of the reaction.

Digitized by Google

A mathematical theory of Natural and Artificial Selection. Part IV. By Mr J. B. S. HALDANE, Trinity College.

[*Received* 11 November, *read* 22 November 1926.]

In such organisms as annual plants, in which successive generations do not overlap, the composition of the $n+1$th generation can be calculated from that of the nth, and the resulting finite difference equation investigated. Where generations overlap we may obtain a similar relation between the compositions of the population at times t and t', but the finite difference equation is now represented by an integral equation. This fact was first pointed out in 1910 by Mr H. T. J. Norton of Trinity College. At a much later date I arrived at the same conclusion, and Mr Norton showed me his results in 1922, stating that he would publish them shortly. He has been prevented from doing so by illness, and, although I believe that all the results here given were reached by me independently, there can be no question that Mr Norton had obtained many of them previously, and had treated the problem rigorously, which I have not done.

The only case considered here is the very simple one in which the intensity of selection is independent of the size of the population. A preliminary lemma will first be discussed.

The growth of a population.

If the death-rate and birth-rate of a population are not functions of its density, its number at any time may be calculated as follows:

Let $N(t)$ be the number at time t. Only the female sex need be considered if the sexes are separate.

$S(x)$ be the probability of an individual surviving to the age x.

$U(t)\,\delta t$ be the number of individuals produced between times t and $t+\delta t$.

$K(x)\,\delta x$ be the probability of an individual between the ages x and $x+\delta x$ producing one (female) offspring. All individuals of this age, both alive and dead, are considered, so that if $P(x)$ be the corresponding function for living individuals only, $K(x) = S(x)\,P(x)$. Then

$$\left.\begin{aligned} N(t) &= \int_0^\infty U(t-x)\,S(x)\,dx \\ U(t) &= \int_0^\infty U(t-x)\,K(x)\,dx \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots (1{\cdot}0).$$

Digitized by Google

Instead of infinity any upper limit exceeding the maximum life of the organism may be taken. Equation (1·0) has been considered by Herglotz*. Let $U(t) = ce^{zt}$. Then

$$\int_0^\infty e^{-zx} K(x)\,dx = 1, \text{ or } \int_0^a e^{-zx} K(x) = 1 \quad \ldots\ldots\ldots (1{\cdot}1),$$

where a is sufficiently large. Since $K(x)$ is always real and zero or positive, the above integral is a monotone function of z when z is real, and can have any real positive value. Hence it has one and only one real root for z, say α_0. The complex roots clearly occur in pairs $\alpha_r \pm i\beta_r$. Then

$$\int_0^\infty e^{-\alpha_r x} K(x)\,dx > \int_0^\infty e^{-(\alpha_r + i\beta_r)x} K(x)\,dx$$
$$= 1.$$

Therefore $$\alpha_r < \alpha_0.$$

If any two functions of t are solutions of (1·0) so is their sum. Therefore

$$\left.\begin{aligned} U(t) &= a_0 e^{\alpha_0 t} + \sum_{r=1}^{\infty} a_r e^{\alpha_r t} \cos \beta_r (t - b_r) \\ N(t) &= c_0 e^{\alpha_0 t} + \sum_{r=1}^{\infty} c_r e^{\alpha_r t} \cos \beta_r (t - d_r) \end{aligned}\right\} \quad \ldots\ldots (1{\cdot}2).$$

In general there will be an infinite number of terms. The values of a_r and b_r, and hence of c_r and d_r, depend on the initial conditions. Where multiple roots occur there will be periodic terms including powers of t as factors. Since $\alpha_r < \alpha_0$, all the periodic terms become negligible compared with the first after the lapse of a sufficient time. That is to say, oscillations of the population about an exponential function of the time are either damped or at least increase less rapidly than the population itself. In particular, if $\int_0^\infty K(x)\,dx = 1$, so that the population is in equilibrium, oscillations are damped and the equilibrium is stable. We are therefore justified in neglecting periodic terms in the solution of equations which only differ by small terms from (1·0), and which occur in the subsequent analysis. It is proposed to discuss the stability of the equilibrium when $S(x)$ and $K(x)$ depend on the number of the population in a subsequent paper.

Incidentally, if $C(x)\,\delta x$ be the probability of any member of the population being between the ages x and $x + \delta x$, then

$$C(x) = \frac{U(t-x)\,S(x)}{N(t)}.$$

* Herglotz, *Math. Ann.* 65, p. 87.

Digitized by Google

Hence, when oscillations have died down,

$$C(x) = \frac{e^{-a_0 x} S(x)}{\int_0^\infty e^{-a_0 x} S(x)\, dx}.$$

This constitutes a new proof of Lotka's* theorem on the stability of the normal age distribution.

Selection of an autosomal factor.

Consider a population consisting, at time t, of $D(t)$ female zygotes possessing a dominant factor A, $R(t)$ female recessives. The sex-ratio at birth is taken as fixed.

Let $F(t)\,\delta t$ be the number of fertile A ova produced between times t and $t+\delta t$.

$f(t)\,\delta t$ be the number of fertile a ova produced between times t and $t+\delta t$.

$M(t)\,\delta t$ be the number of A spermatozoa produced between times t and $t+\delta t$.

$m(t)\,\delta t$ be the number of a spermatozoa produced between times t and $t+\delta t$.

$S(x)$ be the probability of a female dominant reaching the age x.

$s(x)$ be the probability of a female recessive reaching the age x.

$K(x)\,\delta x$ be the probability of a female dominant (alive or dead, as above) producing a female offspring between the ages x and $x+\delta x$.

$[K(x)-k(x)]\,\delta x$ be the same probability for a female recessive.

$L(x)\,\delta x$ be the same probability for a male dominant.

$[L(x)-l(x)]\,\delta x$ be the same probability for a male recessive.

$$S = \int_0^\infty S(x)\, dx, \quad s = \int_0^\infty s(x)\, dx.$$

$$K = \int_0^\infty K(x)\, dx, \quad K' = \int_0^\infty xK(x)\, dx, \quad k = \int_0^\infty k(x)\, dx.$$

$$L = \int_0^\infty L(x)\, dx, \quad L' = \int_0^\infty xL(x)\, dx, \quad l = \int_0^\infty l(x)\, dx.$$

In general the functions $S(x)$, $s(x)$, $K(x)$, etc., will not be functions of age alone, but of $D(t)$, $R(t)$, etc. We make the assumption however that selection and population growth are proceeding so slowly that $k(x)$ and $l(x)$, and $K-1$ are small, and $S(x)$, etc., do not vary appreciably in the course of a generation.

* Lotka, *Proc. Nat. Ac. Sci.* 8, p. 339, 1922.

Digitized by Google

If mating be at random, the rates of production of the three female phenotypes at time t are

$$AA, \frac{F(t)\,M(t)}{M(t)+m(t)}; \quad Aa, \frac{F(t)\,m(t)+f(t)\,M(t)}{M(t)+m(t)}; \quad aa, \frac{f(t)\,m(t)}{M(t)+m(t)}.$$

The group aged x at time t was hatched or born at time $t-x$. Therefore

$$\left.\begin{aligned}
F(t) &= \tfrac{1}{2}\int_0^\infty \frac{2F(t-x)\,M(t-x)+F(t-x)\,m(t-x)+f(t-x)\,M(t-x)}{M(t-x)+m(t-x)}\,K(x)\,dx\\
f(t) &= \tfrac{1}{2}\int_0^\infty \frac{F(t-x)\,m(t-x)+f(t-x)\,M(t-x)+2f(t-x)\,m(t-x)}{M(t-x)+m(t-x)}\,K(x)\,dx\\
&\qquad -\int_0^\infty \frac{f(t-x)\,m(t-x)\,k(x)\,dx}{M(t-x)+m(t-x)}\\
M(t) &= \tfrac{1}{2}\int_0^\infty \frac{2F(t-x)\,M(t-x)+F(t-x)\,m(t-x)+f(t-x)\,M(t-x)}{M(t-x)+m(t-x)}\,L(x)\,dx\\
m(t) &= \tfrac{1}{2}\int_0^\infty \frac{F(t-x)\,m(t-x)+f(t-x)\,M(t-x)+2f(t-x)\,m(t-x)}{M(t-x)+m(t-x)}\,L(x)\,dx\\
&\qquad -\int_0^\infty \frac{f(t-x)\,m(t-x)\,l(x)\,dx}{M(t-x)+m(t-x)}
\end{aligned}\right\} \quad \ldots(2{\cdot}0).$$

Since selection and population growth are slow, we may put $F(t-x)=F(t)-xF'(t)$, etc., $M(t)=\lambda F(t)$, $m(t)=\lambda f(t)$, all to the first order of small quantities. Hence, to this degree of approximation,

$$\begin{aligned}
F(t) &= \tfrac{1}{2}KF(t)+\frac{\frac{1}{2}K[F(t)+f(t)]\,M(t)}{M(t)+m(t)}-\tfrac{1}{2}K'F'(t)-\frac{\frac{1}{2}K'[F'(t)+f'(t)]\,M(t)}{M(t)+m(t)}\\
&\qquad -\frac{\frac{1}{2}K'[F(t)+f(t)]\,M'(t)\,m(t)-M(t)\,m'(t)}{[M(t)+m(t)]^2}\\
&= \tfrac{1}{2}KF(t)+\frac{\frac{1}{2}K[F(t)+f(t)]\,M(t)}{M(t)+m(t)}-K'F'(t).
\end{aligned}$$

Similarly

$$f(t)=\tfrac{1}{2}Kf(t)+\frac{\frac{1}{2}K[F(t)+f(t)]\,m(t)}{M(t)+m(t)}-K'f(t)-\frac{k[f(t)]^2}{F(t)+f(t)}$$

$$M(t)=\tfrac{1}{2}LF(t)+\frac{\frac{1}{2}L[F(t)+f(t)]\,M(t)}{M(t)+m(t)}-L'F'(t)$$

$$m(t)=\tfrac{1}{2}Lf(t)+\frac{\frac{1}{2}L[F(t)+f(t)]\,m(t)}{M(t)+m(t)}-L'f'(t)-\frac{l[f(t)]^2}{F(t)+f(t)},$$

all approximately. Therefore

$$\begin{aligned}
\frac{M(t)}{m(t)} &= \frac{\left[2LF(t)+Lf(t)-2L'F(t)\right]\frac{M(t)}{m(t)}+LF(t)-2L'F'(t)}{\left[LF(t)-2L'f'(t)-\frac{2l\{f(t)\}^2}{F(t)+f(t)}\right]\frac{M(t)}{m(t)}+LF(t)+2Lf(t)-2L'f'(t)-\frac{2l[f(t)]^2}{F(t)+f(t)}}\\
&= \frac{L[2F(t)+f(t)]\frac{M(t)}{m(t)}+LF(t)-\frac{2L'F'(t)[F(t)+f(t)]}{f(t)}}{Lf(t)\frac{M(t)}{m(t)}+LF(t)+2Lf(t)-\frac{2L'f'(t)[F(t)+f(t)]}{f(t)}-2lf(t)}
\end{aligned}$$

Digitized by Google

pproximately,

$$=\frac{F(t)}{f(t)}+\frac{2L'[F(t)f'(t)-F'(t)f(t)]}{L[F(t)+f(t)]}+\frac{lF(t)[f(t)]^2}{L[F(t)+f(t)]^2},$$

pproximately, by solving the quadratic. Therefore

$$\left.\begin{aligned}(K-1)F(t)-K'F'(t)+\frac{L'[F(t)f'(t)-F'(t)f(t)]}{L[F(t)+f(t)]}+\frac{lF(t)[f(t)]^2}{L[F(t)+f(t)]^2}=0\\(K-1)f(t)-K'f'(t)-\frac{L'[F(t)f'(t)-F'(t)f(t)]}{L[F(t)+f(t)]}-\frac{lF(t)[f(t)]^2}{L[F(t)+f(t)]^2}-\frac{k[f(t)]^2}{F(t)+f(t)}=0\end{aligned}\right\} \quad \text{.....(2·1).}$$

If $u(t)=F(t)/f(t)$,

$$\frac{d}{dt}u(t)=\frac{(l+kL)u(t)}{(L'+K'L)[1+u(t)]} \quad \text{.....(2·2).}$$

This is equivalent to equation (2·1) of Part I*, $\frac{l+kL}{L'+K'L}$ being the coefficient of selection. In general this quantity is not independent of t, hence the equation cannot be integrated, but if its upper and lower limits are known, the march of the composition of the population can be roughly calculated from equation (2·3) of Part I. If, however, the population is very nearly in equilibrium, and either dominants or recessives are very rare, more accurate results are possible.

When recessives are rare, $F(x)$ is large and equal to a constant F, $D(x)$ being also large and equal to a constant N. $K=1$, and $F'(t)$ is negligible, while $f(t)$ is small. Therefore

$$K'f'(t)+\frac{L'}{L}f'(t)+\left(\frac{l+kL}{L}\right)\frac{[f(t)]^2}{F}=0,$$

and

$$f(t)=\frac{(L'+K'L)F}{(l+kL)(t-t_0)},$$

where t_0 is an integration constant. But

$$N=D(t)=\int_0^\infty F(t-x)S(x)\,dx=FS,$$

$$R(t)=\int_0^\infty\frac{[f(t-x)]^2 s(x)\,dx}{F(t-x)+f(t-x)}=\frac{(L'+K'L)^2Fs}{(l+kL)^2(t-t_0)^2}.$$

$$R(t)=\frac{(L'+K'L)^2sN}{(l+kL)^2S(t-t_0)^2} \quad \text{.....(2·3).}$$

Hence selection proceeds at the same rate as when generations are separate, with a selection coefficient equal to

$$\left(\frac{l+kL}{L'+K'L}\right)\left(\frac{S}{s}\right)^{\frac{1}{2}}.$$

* Haldane, *Trans. Camb. Phil. Soc.* 23, p. 19, 1924.

Digitized by Google

When dominants are rare, $f(x)$ is large and equal to a constant f, $K = 1 + k$, and $F(x)$ is small. Therefore

$$\left(k + \frac{l}{L}\right) F(t) + \left(K' + \frac{L'}{L}\right) F'(t) = 0,$$

$$F(t) = e^{\frac{(l+kL)(c-t)}{L'+K'L}},$$

where c is a constant of integration. Therefore

$$D(t) = \int_0^\infty 2F(t-x)\, S(x)\, dx = 2SF(t)$$

$$= e^{\frac{(l+kL)(t_0-t)}{L'+K'L}} \qquad \text{.....(2·4),}$$

where t_0 is an arbitrary constant. Selection therefore proceeds as when generations are separate, but with a selection coefficient

$$\frac{l+kL}{L'+K'L}.$$

When the death rates and fertility rates are the same in the two sexes, or in a hermaphrodite species, we have in general,

$$\frac{d}{dt} u(t) = \frac{ku(t)}{K'[1+u(t)]} \qquad \text{.....(2·5),}$$

when recessives are rare,

$$R(t) = \frac{K'^2 sN}{k^2 S(t-t_0)^2} \qquad \text{.....(2·6),}$$

when dominants are rare,

$$D(t) = e^{\frac{k(t_0-t)}{K'}} \qquad \text{.....(2·7).}$$

Now k has approximately the same meaning as when generations are separate, provided $K = 1$. Hence the two cases become comparable if we choose our unit of time, or "generation," so as to make K' or $\int_0^\infty xK(x)\, dx = 1$, as in the calculations of Dublin and Lotka* on the rate of increase of a population. In each case, if functions analogous to $S(x)$, $s(x)$ are known for the males, their numbers can be calculated.

* Dublin and Lotka, *Journ. Amer. Stat. Assoc.* 1925, p. 306.

Digitized by Google

Selection of a sex-linked factor.

Here, using the same notation as above, we find

$$\left.\begin{aligned}(K-1)F(t)-K'F'(t)+\frac{L'[F(t)f'(t)]-F'(t)f(t)]+lF(t)f(t)}{2L[F(t)+f(t)]}=0\\(K-1)f(t)-K'f'(t)-\frac{L'[F(t)f'(t)-F'(t)f(t)]+lF(t)f(t)}{2L[F(t)+f(t)]}-\frac{k[f(t)]^2}{F(t)+f(t)}=0\end{aligned}\right\}\ldots(3{\cdot}1).$$

In general, if $u=F(t)/f(t)$,

$$(L'+2K'L)\frac{d}{dt}u(t)=lu(t)+\frac{2kLu(t)}{1+u(t)} \quad \ldots\ldots(3{\cdot}2),$$

an equation analogous to (2·0) of Part III*.

When recessives are rare,

$$(L'+2K'L)f'(t)+lf(t)+\frac{2kL[f(t)]^2}{F}=0,$$

where F is the birth-rate of dominants. Three cases occur:

(*a*) If kL is negligible compared with lN, which will be the case if selection is of the same order of intensity in the two sexes, or more intense among males, then

$$f(t)=e^{\frac{l(t_0-t)}{L'+2K'L}} \quad \ldots\ldots(3{\cdot}31).$$

The number of recessive males is proportional to $f(t)$, of recessive females to its square.

(*b*) If kL is of the same order of magnitude as lF, then

$$\frac{2kLf(t)}{2kLf(t)+lF}=e^{\frac{l(t_0-t)}{L'+2k'L}} \quad \ldots\ldots(3{\cdot}32).$$

Hence if V, v are quantities corresponding to S, s for the male sex, the proportion of recessive males is

$$\frac{lv}{2kLV\left(e^{\frac{l(t-t_0)}{L'+2K'L}}-1\right)},$$

of recessive females

$$\frac{l^2s}{4k^2L^2S\left(e^{\frac{l(t-t_0)}{L'+2K'L}}-1\right)^2}.$$

* Haldane, *Proc. Camb. Phil. Soc.* 23, p. 363, 1926.

Digitized by Google

(c) If kL is much larger than lF, i.e. selection is confined to females, then

$$f(t) = \frac{(2K'L + L')F}{2kL(t - t_0)} \quad \text{.....(3·33).}$$

Hence the proportion of recessive males is

$$\frac{(L' + 2K'L)v}{2kLV(t - t_0)},$$

of recessive females

$$\frac{(L' + 2K'L)^2 s}{4k^2L^2S(t - t_0)^2}.$$

When dominants are rare,

$$F(t) = e^{\frac{(l+2kL)(t-t_0)}{L'+2K'L}} \quad \text{.....(3·4).}$$

The number of male dominants is proportional to $F(t)$, that of females being double that of males.

When the intensity of selection is equal in both sexes, these equations simplify to

$$\frac{d}{dt}u(t) = \frac{ku(t)[3 + u(t)]}{3K'[1 + u(t)]} \quad \text{.....(3·5).}$$

$$f(t) = e^{\frac{k(t_0 - t)}{3K'}} \quad \text{.....(3·6),}$$

$$F(t) = e^{\frac{k(t_0 - t)}{K'}} \quad \text{.....(3·7),}$$

analogous to equation (7·2) of Part I.

Discussion.

The most satisfactory table of $K(x)$ known to me is that given by Dublin and Lotka* for certain American women. Here the population is growing, and $K = 1{\cdot}17$, while $\frac{K'}{K}$, the length of a "generation," is 28·45 years. No satisfactory values of k are known in the present state of genetics, though the data on mice discussed in Part I suggest that here $k = {\cdot}04$ approximately. In man mating is highly assortative for age, and the above formulae cannot be applied. Moreover, a change in the coefficient of correlation between the ages of spouses would undoubtedly affect the values of $K(x)$, etc., if other conditions remained equal. Thus old men would beget more children if they were more likely to have young

* Dublin and Lotka, *loc. cit.*

Digitized by Google

wives. It is thus impossible to calculate the effect of this correlation on selection. But a consideration of the extreme case when the age of the wife fixes that of the husband makes it clear that selection must follow equations of the type here arrived at, with changes in the parameters only.

Summary.

Expressions are found for the progress of slow selection in a Mendelian population where generations overlap. The changes are very similar to those which occur when generations are separate.

Digitized by Google

Digitized by Google

PROCEEDINGS
OF THE
Cambridge Philosophical Society.

What determines the Resistance and the Tilt of an Aeroplane? By Sir JOSEPH LARMOR, Lucasian Professor.

[*Read* 31 January, *received* 1 February 1927.]

It is recognized that the main hope for advance in a rational theory of aeroplane propulsion lies in extensive use of the general principle of conservation of resultant momentum, even more than on considerations of energy; for energy can be dissipated by friction, while the relation of momentum to force is absolute.

For a field of vortices in unlimited fluid, without solid boundaries, remarkable general kinematical expressions for resultant hydrodynamic linear momentum (P, Q, R) and angular momentum (L, M, N) with respect to a selected origin, and for the kinetic energy (T), the former pair involving the distribution of vorticity (ξ, η, ζ) alone, were given long ago by J. J. Thomson in an Adams Prize Essay on *Vortex Motion* (1883, pp. 5–8). They appear, at any rate on first view, to be precisely adaptable to application of the kind above indicated: and they have the advantage moreover of being in three dimensions, dispensing to some degree with the usual limitation of theoretical discussions to plane motions.

The formulae will now be quoted *. They are of types, $\delta\tau$ being an element of volume, and there being either no boundaries or else the suitable vortex sheets spread over the surfaces of the immersed solids,

$$P = \rho \int (y\zeta - z\eta)\, d\tau, \qquad L = -\rho \int (y^2 + z^2)\, \xi\, d\tau,$$
$$T = 2\rho \int \{u(y\zeta - z\eta) + v(z\xi - x\zeta) + w(x\eta - y\xi)\}\, d\tau.$$

It is easier to verify these results than to discover them: in so doing we shall extend them so as to apply to a part of the field of motion instead of the whole. Thus, on substituting †

$$(\xi, \eta, \zeta) = \frac{1}{2}\left(\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z}, \frac{\partial u}{\partial z} - \frac{\partial w}{\partial x}, \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right)$$

and integrating by parts, with some easy reduction, we obtain for these integrals taken over any bounded region of fluid, itself of any viscous quality, regular or not if only it be of constant uniform density, the values

$$P = \tfrac{1}{2}\rho \int (-yum - zun + yvl + zwl)\, dS + \int \rho u\, d\tau,$$
$$L = -\tfrac{1}{2}\rho \int (wm - vn)(y^2 + z^2)\, dS + \int \rho\, (yw - zv)\, d\tau,$$
$$T = \tfrac{1}{2}\rho \int 2\, \{(lu + mv + uw)(xu + yv + zw)$$
$$- (lx + my + nz)(u^2 + v^2 + w^2)\}\, dS + \int \tfrac{1}{2}\rho\, (u^2 + v^2 + w^2)\, d\tau.$$

* The sign of (L, M, N) has been reversed.

† The factor $\frac{1}{2}$ in this definition identifies vorticity with differential rotation. The formulae imply that the velocity falls off towards infinity owing to all the vortical sources being at finite distance.

Digitized by Google

The second terms in these values of the integrals P, L, T are, as anticipated, the momenta and the energy in the region, the first terms being surface integrals over its boundary, which is supposed simply connected unless the motion is acyclic.

For any region of the fluid the content as regards momentum-components and energy, which will be denoted for it by accented letters, are thus given in terms of the volume integrals P, L, T. expressed above, and of boundary integrals as transferred across to the other sides of these equations. The boundary integrals do not contribute practically towards the forces, as they merely fluctuate *in steady motion*. Indeed if the boundaries are fixed, the surface integral in the rotational momentum round the origin, in the present form, vanishes. The essential vector character of local momentum may be brought out by a slightly modified formulation of the result. We have, for any region, now in explicit invariant form, r^2 being $x^2+y^2+z^2$, U denoting velocity, U_n and U_τ its component transverse to the boundary and along r, and the bracket $|\ldots|$ representing a scalar product, the momentum and energy

$$P' = \rho \int (y\zeta - z\eta)\, d\tau + \tfrac{1}{2}\rho \int \{u\,|r\,dS| - |Ur|\,l\,dS\},$$

$$L' = -\rho \int (x^2+y^2+z^2)\,\xi\, d\tau + \tfrac{1}{2}\rho \int r^2 (wm - vn)\, dS,$$

$$T' = -\rho \int \begin{vmatrix} x & y & z \\ u & v & w \\ \xi & \eta & \zeta \end{vmatrix} d\tau - \rho \int U_n U_\tau\, r\, dS + \rho \int U^2 |r\, dS\,.$$

when vorticity is defined as velocity of spin, as above. It thus appears that the resultant translatory momentum is the same as would be derived from endowing each element of fluid with an ideal translatory momentum equal per unit mass to the moment of its vorticity round the origin; this part for the whole field of motion involves no resultant torque round the origin, but it is supplemented by a moment of momentum with ideal contribution from each element equal per unit mass to its vorticity multiplied by the square of its distance. The significant feature is that only the vortical regions of the field of fluid motion, the current whirls, are involved in these expressions. The kinetic energy would be derived similarly by taking from each vortical part of the motion an ideal contribution equal per unit mass to the determinantal product of the distance, the velocity, and the vorticity.

For instance, a straight endless horizontal vortex of strength ϖ would have a lifting force per unit length $\rho \varpi v$, where v is its horizontal velocity, which is transferred to distance $2v$ from the origin by the accompanying torque, a statement which acquires sense when there are compensating return vortices. For an aerofoil, as will appear, the completion is by a return sheet of vortices above.

Digitized by Google

These formulae are moreover of universal application, being merely kinematical transformations of the actual distributions of momentum and energy*. The governing dynamical feature that makes them useful now enters in the circumstance that, outside an enclosing boundary far enough away from the local irregular motions near an aeroplane, the movement of the air may be taken to be on the whole differentially irrotational, any rotational character, such as alone enters into the integrals expressing the momentum, being restricted in the main to the substance of vortex rings associated with the wake that it leaves behind†. It may be noted that the momenta provide an ideal means of descriptive dynamical exploration, as the contributions from different sources are strictly superposable. The confused momentum inside the interface, with which fortunately we are not directly concerned, as it is steady, would have to include terms involving the slip-surface or vortex sheet over the boundaries of the solids.

The genesis of these vortices shed off from the moving body, as traced to the development of unstable surfaces of slip in the fluid, was explored by Helmholtz himself, the great discoverer of the laws of vortex motion, was then developed by Kirchhoff, Kelvin, Rayleigh, and has been improved into a general analytic method by J. H. Michell, Love, Lamb and others. The relation of this slip-theory, so far as it may be accepted, to the phenomena of resistance and support is intimate and fundamental, especially when it is recognized as here that the interchanges of momentum in the field of fluid motion may be associated wholly with the trail of vortex whirls that are shed away from the propelled machine [or carried along with it].

Let us suppose that a body like an aerofoil, with a blunt front and rather sharply curved trailing edge, is travelling through air, or, more conveniently for the present argument, is anchored as in the diagram in a stream of air. The problem set by Helmholtz was to determine a mode of flow, of type suggested by observation as *infra*, involving a sheet of slip as indicated in the diagram, such as would ease off the infinities at sharp edges that occur in the continuous theoretical solutions for motion of frictionless fluid. Such a sheet of slip does in actual flow arise, but is soon broken up as it has obviously no stability: it is constantly being renewed

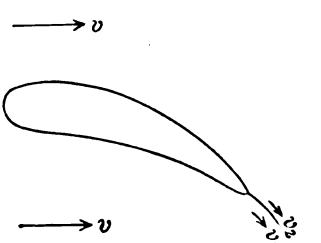


Fig. 1.

* Applied to a solid of density ρ_0 moving in water of density ρ, the integrations would naturally be taken over all space, at the constant density ρ, leaving over a separate momentum of the solid at reduced (Archimedean) density $\rho_0 - \rho$.

[† or carries with it in renewed circulation, as *infra* at end.]

Digitized by Google

close to the edge, but as the fluid moves away carrying it along, with blurring by diffusion, increasing corrugation of the unstable form breaks the sheet of slip into isolated fragments; these, having thus acquired the status of very flat independent vortices, soon shrink each of them into the stable circular form of cross-section. This course of events is presumably what is meant by the phrase used by Helmholtz and his successors, that a surface of slip 'curls up' into a row of vortices*.

It was the practical study of vortices, recognized obviously, in advance of theory, as for some dynamical reason a sub-permanent feature in actual fluid motion, that led Helmholtz into one of the most brilliant chapters in modern theoretical physics. His original problem was the slip-sheets of the air-jet, ultimately breaking into vortices, formed at the lip of a blown organ-pipe, which by their rhythmic succession establish and also maintain regular air-vibrations with periods that are imposed within the pipe itself: thus providing, outside the acoustic application, an arresting natural dynamical model for general physical rhythm, even possibly quantified, whose potentialities are still far from being exhausted. Helmholtz traced the initiation of vortices in fluid to actual surfaces of slip. But such sharp slip, if it really occurs, being a clean tear in the fluid, stands itself even more in need of elucidation. Lord Kelvin at one time ascribed abrupt slip to an initial cavitation† where the fluid flows round a sharp bend with resulting high local speed and thus greatly diminished pressure: such a cavity, drawn out presumably along the flow and continually renewed, would on ultimately closing down provide an adequate cause, initially at any rate, for a sheet of perfect slip. Long after that time, the occurrence of cavitation, leading perhaps by its degradation through such slip-sheets to local whirl of the fluid as a whole along with the screw, imposed conspicuous speed-limits on some of Sir Charles Parsons' practical problems of propulsion by means of turbines.

Circuits made up of the same particles of fluid remain ordinarily

* More generally, familiar observation appears to indicate that laminar motion, in a layer that would be unstable by the Rayleigh rule, would be stabilized by inserting in it a row of vortices: cf. a theoretical note by Kelvin, *Scientific Papers*, IV, p. 186.

† *Scientific Papers*, vol. IV, Hydrodynamics; a playful epistolary controversy between Stokes and Kelvin on the possibility of slip of the Helmholtz type is there abstracted. The force of the explanation by cavitation is substantially reduced for the case of gases such as air. It is also asserted from observation that vortices break right off from the back of the edge.

Incipient cavitations, continually collapsing against the wing itself, might very well maintain an irrotational circulation round it. The loads actually sustained are astonishing. In any case the momentum criterion involves that a free wing has to adjust its tilt so that the double vortex sheet left behind it is at its source of small breadth and so of slight account.

Eddy resistance, in relation to fineness of lines in the rear, was considered by Stokes as early as 1843: cf. *Collected Papers*, vol. I, pp. 53, 99, 311.

Digitized by Google

complete: no vortex filaments can cut across them, and the circulations in them remain unchanged whether they traverse vortices or not. The cavitation is what opens a breach in such circuits, permitting slip which becomes vortex motion to pass across, thus introducing abrupt change into their circulations.

This local discontinuity from velocity v_1 to v_2 in crossing the slip-surface, which after travelling a short distance thus breaks up into whirls, must establish circulation in the smooth irrotational flow of the fluid as a whole around the front parts of the initiating body. For the velocity of relative slip is a discontinuity in the value of the tangential gradient of the velocity potential of smooth flow, arising in crossing the sheet: therefore, owing to accumulation of gradient along the sheet, that potential is itself finitely discontinuous in crossing the slip: and the line integral of its gradient in a path around the body, that is, the circulation, as Kelvin named it, instead of being null, as it would be in a complete circuit if the potential were acyclic and so came back to its initial value, represents a cyclic flow determined mainly, as a problem of irrotational motion of a different and discontinuous type, as set by Helmholtz, by the general shape of the body, except in the parts close to the slip-surface, where the circulation could be imagined as adjusted for adjacent paths by superposed local disturbance not much affecting the motion at a distance.

The recognition and analysis of this secondary circulation round the body has, as appears, been utilized by Lanchester, Prandtl and others, as a practical mode of exploring the character of the supporting pressure on aerofoils, this being greatest underneath, where in the relative steady problem velocity of the fluid as held back by the circulation is least. It presents itself as additional to direct exploration in detail of the effect of the vortex wake: the wake-effect being itself transmitted as in solutions of the Helmholtz slip-problem into a definite modification of the frontal motion with consequent change of pressure along the wing, which may be smooth enough over the greater part of the surface for treatment by standard hydrodynamic methods.

In the analytical solutions of the slip-problem, it appears that abruptness of change in the motion near the sharp edge might still remain excessive, notwithstanding easing by the sheet of slip; if that were so it would even prevent smooth formation of any such sheet. The conditions determining the slip-surface have therefore recently, as it appears, been restricted further by Joukowski, by the postulate that near the edge the velocity is to keep within practicable bounds suitable for unbroken flow, or in the abstract mathematical formulation is to remain finite. And this mode of fixing the value of the secondary circulation round the body seems to have had practical application.

Digitized by Google

But from our present point of view the formula for (L', M', N'), which determines the rotational momentum coming within a region, appears to intervene. Referring to the centre of mass of the aeroplane as origin of coordinates, each vortex in crossing our arbitrary boundary separating the turbulent motion near the aeroplane from the more steady flow outside, carries away momentum from the system inside; and the formula shows that the moment of such momentum continually increases as the vortex recedes from the origin by a factor r^2, unless a proper balance of vorticity positive and negative is established. Such continued abstraction of moment of momentum, from the inner turbulent system around the aeroplane, would have its reaction on it in the form of a torque increasing as r and tending to rotate the aerofoil. But all the vortices formed along the same slip-trail have the same direction of spin, that of the vortex-sheet from which they are developed: they would thus provide torques all conspiring together, so that the aeroplane, if free, would at once turn round by reaction and adjust its presentation to the stream so as to get rid of their torque,—or more precisely, so as to obtain equilibrium between the pull of the screw, the weight of the aeroplane and the lift-force and the drag of the wake.

The conclusion appears to be involved, that an aerofoil free to alter its tilt and speed cannot originate permanently a single trail of vortices, for that would produce mainly a torque to turn it round. It must produce a double trail, one sheet starting, as actually appears, from near the edge, the other from the back surface of the foil in its neighbourhood, 'dead water' being more or less fully developed between these trails with their two sheets of vortices thus spinning in opposite directions. As the speed of the aeroplane is altered, the pull of this wake would thus have to adjust itself by change of tilt into a single force passing through the intersection of the pull of the screw and the line of weight of the machine, which might then be taken for origin in the formulae.

The condition that determines the tilt is that the two sheets of vortices should cancel each other as regards moment of momentum about this point of intersection; which requires that the aggregates of their strengths should be everywhere nearly equal and opposed in order to annul L. This consideration of L null would be the criterion fixing the value of the secondary circulation of the Lanchester-Prandtl discussions: the slips on the sheets would not be exactly equal and opposite, else it would vanish*. It would appear, on the present view, in case the opposite vortices in the two sheets were not far from being equal and opposite, that at first at any rate they should be about equidistant in pairs from

* Recent experiments by Bénard, reported in *Comptes Rendus*, seem to show that regular spacing on the two sheets, which simplifies the Karman mathematical theory by making the motion steady, is not essential for stability of the vortex trail.

Digitized by Google

the centre of mass of the aeroplane, and therefore situated some way above it. If the conjugate straight vortices are thus imagined as bracketed in pairs, in the order of their origination, each pair if of exactly opposite strengths would be analogous to one elongated vortex ring: indeed the vorticity must actually curl round at the ends like a flattened ring. The momentum of a ring is transverse to its plane and acts at its centre of form: as being above the mass-centre of the machine it must here be downwards in order that the lifting reaction may be upwards. This is consistent with two sheets of vortices alternately spaced, each edge vortex below, with its determined direction of spin conditioned by the edge trending downward, being regarded as associated with a next following shoulder vortex. Vortices of opposite spin tend to attract each other. Friction in the wake will waste energy, but cannot alter momentum as a whole.

The relevant considerations may thus be presented on a less abstract basis by formulating the momentum associated with a single isolated thin vortex ring. If ϖ is the total strength integrated over its small cross-section, which is necessarily constant along its circuit, the formula for the momentum (P, Q, R) of its field reduces to a line integration round its circuit, of type

$$P = \rho\varpi \int (y\,dz - z\,dy):$$

thus the component of the momentum in any direction is $\rho\varpi$ multiplied by twice the projection of the aperture of the ring on a plane transverse to that direction.

It may be remarked that, on these principles, a Karman double trail of straight vortices, completed into elongated rings, each of strength ϖ, spread alternately at intervals a on each of the sheets whose distance apart is h, N conjugate pairs being emitted per unit time, ought to exert per unit breadth a lifting force $\rho a\varpi N$, a drag $2\rho h\varpi N$ and a torque $\rho\varpi N\,|r^2 - r'^2|$, where the last factor is the difference between the initial and the ultimate values of the difference of squares of their distances from the selected origin. They carry away energy $4\rho h\varpi Nv$ per unit time, which would give a drag twice too great, were it not that their own velocity v is only half of Na on account of the dead water between them.

The resultant momentum associated with a vortex ring is in direction such that the projected area A on a plane transverse to it is maximum, and is equal to $2\rho\varpi A$: the ring-system must travel along this direction.

The resultant associated momentum for any travelling vortical system however complex must be expressible canonically as a linear momentum Λ round a definite central axis and a moment of momentum Γ, or angular momentum in the terminology introduced by Thomson and Tait, around that axis.

Digitized by Google

When a single vortex is isolated, it settles into circular form, and this central axis must be transverse to its plane through its centre. Then Λ is $2\rho\varpi A$, where A is the area of the circle, while Γ must be by symmetry null. If however the vortex is a compound one, for example, of the type made up of vortex filaments interlaced loosely in spirals, wound round each other like a rope but without contact, after the manner of vortex-atom models long ago investigated by J. J. Thomson, the stranding may be dextral or laeval, and Γ can exist.

Any such complicated interlaced vortex system must have a central screw-axis of momentum. Its system of momentum must be conserved as its form travels through the fluid, either with vacuous core or with a core of fluid convected along with it; therefore it must settle down to travel along a straight line, this central screw-axis.

When a travelling group of vortices becomes obliterated by viscosity, its momentum survives. The expression for angular momentum, as it involves a factor r^2 yet cannot increase indefinitely, requires that for an isolated group

$$\rho \int d\tau \, (\xi, \eta, \zeta) = 0,$$

so that the resultant of the component vorticities in any direction is null, as is readily verifiable. It shows also that for a point on the axis of the momentum-torque the components of (ξ, η, ζ) in a plane transverse to the axis each multiplied by the square of its distance from the axis must have null vector resultant: thus revealing the locality of the axis of progress by inspection as a sort of central line of the system.

These considerations may perhaps be further illustrated by the case of a full-shaped body like a symmetrical air-ship regarded as throwing off from its rearward shoulders a succession of ring-vortices, now situated on a single conical sheet, and also minor secondary vortices from edges: the motion would settle down so that this system of vortices travels along their resultant axis, which is nearly that of the body. From this descriptive point of view a rudder deflects the travelling body by a process which may be regarded as the sending away of a stream of vortices from its edges, abstracting momentum in an asymmetric manner which involves a reacting torque on the body including the steady turbulence around it. When the rudder is released the system settles back into the configuration and mode of progress which preserve unaltered the main regular momentum, unless indeed the body has been deflected so far that it swings away towards a new position of stability. Thus any free unsteered body, immersed in water or air, turns flatways to the stream: for bodies such as air-ships, the further classical theory for steering which operates by arresting the be-

Digitized by Google

ginnings of such cumulating instability, as constructed by G. H. Bryan in the early days of aviation, has all along been in full practical application.

The momenta associated immediately with the nearly straight linear elements of a vortex ring are mainly angular momenta around them; but for the complete ring they counteract, and for symmetric forms cancel out. The integrations by parts, by which momentum actually belonging to the inertia of the fluid in the field is attached formally to the vortices, have rearranged its location artificially in the general formulae with which we began.

An ideal flattened air-ship with smooth contours, a long aerofoil however being hardly included even as a limit, may perhaps in suitable conditions be conceived to throw off elliptic vortex rings from its shoulders, which would oscillate through the circular form as they recede and finally rest in that form by virtue of steadying influence of frictional agencies. In no case can a train of rings continue to be shed off obliquely to the motion of the stream, because their axes of momentum, being then inclined to the wake, would acquire greater and greater moment with respect to the source as they recede from it.

The motion of a flat body transversely through fluid can also with careful initiation take a different form, sometimes apparently smooth and stable without much wake, by its acquiring a thick vortex ring following close behind it which eases off the motion round the convexities of its surface.

From the present point of view it becomes more obvious that the support of a banked aeroplane is transverse to the wings, and thus oblique and so adapted to counteract the centrifugal reaction to its motion in a curve.

The mathematical theory of the interaction of a set of vortex rings travelling along a common straight axis, as developed first by Kirchhoff, gives another illustration of this descriptive mode of analysis by momentum. The total momentum must be conserved; and now also the energy, as there is in this ideal problem supposed to be no turbulence to accelerate its dissipation. If the vortices are thin rings, the one of strength ϖ_n having radius r_n and velocity v_n, these conditions lead by the initial formulae to two relations, sustained as these quantities r, v change during the motion, namely

$$2\rho\Sigma\varpi_n\pi r_n^2 = M_o, \text{ constant}, \qquad 4\rho\Sigma\varpi_n\pi r_n^2 v_n = T_o, \text{ constant}.$$

For two vortex rings, continually threading one another, they suffice as is familiar to determine the more prominent features of the motion.

The present order of ideas, to be of utility in practical aeronautics, would have to imply that, as regards aggregate momentum,

Digitized by Google

the turbulent unanalyzable wake of the aerofoil can be neglected in comparison with its explorable large-scale vortex motion. The turbulence is itself mainly irregular vorticity on a minute scale, quickly dissipated by internal friction, and perhaps involving negligible resultant momentum. If so, the steady dynamics of the system, propelled in air, could in theory be formulated in terms of estimation of the momentum it thus leaves behind; but how far that plan can be carried into practice, and the principle so become an observational aid to design, must be a matter for the domain of experience, in which doubtless extensive stores of technical information have already been garnered.

Postscript.—This note, perhaps too venturesome, has been drawn up by the light of nature, aided at the end by hints gained from H. Glauert's recent book on *Aerofoil and Airscrew Theory*. Since it was completed the writer has been able to inspect two memoirs in *Phil. Trans.* 1925, tracing the flow round aerofoil models, by L. W. Bryant and D. H. Williams and by A. Fage and L. F. G. Simmons, which contain a mass of interesting experimental plottings and verifications, that only familiarity with the subject could render assimilable in a short time, supported in part by reference to a large and inaccessible technical literature. It is however here relevant to explore the dynamics of the relation of the presumed definite circulation round the machine to the lift, which can be done on lines indicated in an appendix to one of the memoirs by G. I. Taylor. If the axis of z is taken vertically upward, the rate of change of the vertical momentum content of the aeroplane and the region surrounding it is, when the direction (l, m, n) of the normal to the element of the interface δS is measured inwards, and p represents the fluid-pressure,

$$Z = \int pn\,dS + \rho \int (lu + mv + nw)\, w\,dS$$
$$= \int (p + \tfrac{1}{2}\rho \text{ vel.}^2)\, n\,dS + \rho \int \{luw + mvw + \tfrac{1}{2} n\,(w^2 - u^2 - v^2)\}\, dS.$$

We have here to consider the relative problem of the aeroplane at rest within this fixed boundary and the air flowing through, in order to get steady conditions in which no momentum is used up in changing the motion of the air within the interface, so that Z may be wholly supporting force. Replacing u by $U + u'$ where U is the reversed speed of the aeroplane and (u', v, w) expresses the actual motion of the air, the second term in Z, say of form $f(u, v, w)$, reduces to

$$-\tfrac{1}{2}\rho U^2 \int n\,dS + \rho U \int (lw - nu')\, dS + f(u', v, w).$$

Of this the first term, only however when it has been integrated all over the complete interface, contributes nothing. The second term is ρUC per unit length along the foil, where C is the circulation round the cross-section of the boundary interface: it applies for the

Digitized by Google

middle part of the motion, where it is practically plane; also more roughly elsewhere except at the ends, because the circulation is nearly equal to its projection on the central plane; while beyond the ends the circulation falls away to a null value.

The third term depends on squares of the actual velocity of the air. Near the aerofoil it is far from negligible, but at a distance of several diameters the velocity begins to vary with distance r as r^{-2}, so that the integral varies as r^{-3} and becomes negligible at a moderate distance. Also it is a residual effect between above and below. And the first term in Z, which includes the effect of pressure, vanishes everywhere in the irrotational part of the steady flow. Apparently this statement represents the case that can be made for the dependence of the lift on a circulation around the wing without vorticity.

This formula for lift, named as it appears after Kutta and Joukowski, would hold more precisely, as regards its pressure term, for the slow unbroken motion near a fair-shaped body: though on account of the first term, which disappears only by integrating out, the result does not express the distribution of the supporting force but only its total amount. In actuality the argument is upset in the region of the wake; but not much, if the wake is narrow. The formula, applicable apparently only for the value of C round moderately distant circuits, is also incomplete in that it does not involve the line of action of the supporting force, itself be it noted independent of the tilt to this approximation, and so leaves the tilt unregulated.

It is essential to note that this circulation C remains constant, in a vortex field, only when the circuit continues to be made up of the same particles of fluid: thus as new fluid comes up to the aerofoil it has to be continually renewed*, in fact by creation of new vortices from the trailing edge. On the present view the circulation is continually being generated by the slip-process which sets up the Helmholtz type of irrotational motion all round the foil: at a distance of several diameters it tends to a constant value round all circuits and the motion there is on the average a field of cyclic irrotational motion, such as would take place smoothly round an ideal core of some form probably not that of the foil, the part of the circulation in crossing the wake being thus small and alternating.

We may expose the amount of the practical discrepancy in C more definitely, also on lines carried through in detail by G. I. Taylor, by considering the problem of the relative motion for an air-ship followed by a conical surface of slip having no motion

* A vortex absolutely steady relative to the wing would by the present formulation have constant momentum in the relative motion, and therefore would exert no force, which is in confirmation of this statement.

Digitized by Google

within it. This is permissible notwithstanding instability: for the slip may for theory be regarded as stabilized, within narrow limits, by very slight internal constraint, such as the presence of an ideal internal friction along it would provide. Continuity in the pressure requires that the velocity (v_0) be uniform along the slip-surface, for $\varpi + \frac{1}{2}\rho$ vel.^{2} must be the same on both sides of the sheet of slip while inside it the velocity is now null. The conversion of an enclosing circuit $ACBA$ crossing the slip cone at A and B into another form $A'C'BA'$ crossing it at A' and B subtracts $v_0 AA'$ from the circulation round it. Thus the circulation is not the same for adjacent circuits, but only for those that avoid the wake: for the others it depends on where the circuit strikes and leaves the stagnant relative wake. When the vorticity is now (theoretically) released from the slip-surfaces, and moves away as free vortices into the wake now itself in internal motion, it is the part of the path of the circuit affected by the wake that renders the circulation variable, two such paths making contributions which differ by twice the total vorticity which is enclosed between them. That, by the above, is slight when the paths cross the wake transversely. Actually for a Karman double trail of vortices the circulation fluctuates, as the crossing place is moved outwards along the wake: with an amplitude equal to twice the vorticity of one of the vortices. This implies an effect on the fluid motion due to the influence of compensating adjacent positive and negative vortices, which may perhaps be presumed to be local and so ineffective except close to the foil. The disturbing effect of this vorticity on the flow, after it has settled down to steady conditions, may thus be presumed to be local: it would be specially intense and turbulent near the ends of the aerofoil, where the direction of the vorticity veers round into the reversed or negative part of the elongated closed vortex, corresponding to the rings of the vortex cone following a more rounded body. The result as regards the aerofoil may be nearly the same as if there were no wake, but steady cyclic motion all round the central parts of the foil tending at a considerable distance to the uniform circulation appropriate to irrotational flow, as determined by the considerations here sketched; the supporting force being of this standard value ρUC per unit transverse span, a value which, of necessity in the present train of ideas, is independent not only of the presentation of the aerofoil to the stream but even largely of the actual inner motion as determined by its size and shape.

The theory of lift, drag, and torque, for a body of aeroplane shape, monoplane or biplane, is developed in a very interesting manner, for two-dimensional flow, in terms of strict conformal transformations initiated it seems by v. Mises, in Glauert's book, ch. VII *seq.* The result includes a condition for constancy of

Digitized by Google

torque, which would determine tilt. At first sight the constancy comes out of the Joukowski condition, which has been here put aside on the ground that this constancy could not subsist for a single vortex trail, and that a double trail with opposite vorticities must be required. But it appears that this condition of single slip-surface with velocity finite is introduced merely to give a value for the circulation round the wing: while calculations made roughly from the slip on the surface of a thin nearly flat wing, with a circulation to express the effect of the wake, lead plausibly without any restriction to relations of the same type. The point is perhaps important, as this latter theory, with which the names of Munk, Glauert, and Birnbaum are associated (Glauert, pp. 87–92), is reported to be in close agreement with experimental determinations.

Feb. 1.—It was pointed out when the paper was read that this reference (p. 621) to the Joukowski theory is a misunderstanding. That theory aims at determining a vortical whirl round the wing such as will just *abolish all tendency* to a sheet of slip at the edge*. The proof that this is theoretically possible is certainly a new departure, which can depose the wake from a controlling position for the case of flat bodies moving nearly edgeways. When this course is taken, as seems however justified only for a wing of unlimited length, this field of circulation would travel on accompanying the wing just as the field of circulation of a vortex ring travels on with its core, without need of renewal. It is now however an additional straight vortex, with its circulation: and it adjusts itself to accompany the wing instead of falling away down the wake, there being no wake at all in the ideal case. The theory of momentum here set out would derive a lifting force $\varpi\rho v$ and no drag, where v is now the velocity of the wing, and ϖ is the total vorticity, reckoned as spin, which is equal to half the circulation C round a circuit *far enough away* to be outside any vorticity in the field†. The drag on the wing is on such a view due only to an accidental wake of vortices, now itself really secondary, which also makes a contribution to the lift: and the drag at any rate is reduced in proportion if the breadth of the wake between its compensating positive and negative faces can be adjusted to be small. This application of the momentum formulae replaces, but only for an infinitely long wing, the algebra of this Postscript. It will be observed that there is now no implication that this whirl in which the wing is supposed to become enveloped is a wholly irrotational one sliding by slip rollers at the surface of the ring: what an experimental verification of the relation of

* See L. Bairstow, *Applied Aerodynamics* (1920), p. 360: also L. Prandtl, *Die Naturwissenschaften*, Feb. 1925.

† The discrepancy with the formula for a ring (p. 623) is adjusted by considering the ring made up of two conjugate straight vortices connected by vortical arcs at the ends: the latter make a contribution equal to that of the straight parts.

Digitized by Google

lift to circulation can confirm is that a vortex cylinder becomes established around the wing with *some distribution of vorticity* such as can in the actual circumstances, not two-dimensional, travel steadily along with the wing, and that the circulation C of the formula is to be measured round a circuit far enough away to be outside this vorticity. Approximate verification of the theoretical formula for an irrotational flow would imply that the region of this vorticity clings close to the wing; which would permit the distribution of the lift over the surface of the wing to be determined as above indicated. Anyhow this very substantial easing of the conditions of the problem makes it more intelligible that the instability at a sharp edge can really get itself largely surmounted in this way, even when one remembers that the wing is not actually, as the two-dimensional theory requires, very long compared with its breadth.

[The lesson that is here enforced is that we must deal only with completed vortex rings: thus in two-dimensional analysis each straight vortex must be bracketed with a return vortex, as is familiar for the analogous case of electric currents.

There seems to be a hypothesis which does continue the circulation around the wings by vortices trailing from their outer ends and completing a circuit across the wake. The momentum of such a steady completed vortex travelling in fluid is (p. 623) of the order $2\rho\varpi A$, where A is the area of its circuit; the transfer of this momentum seems to involve no supporting force on the wings, but a torque $2\rho\varpi Av\cos\alpha$, where α is the inclination of its area A to the direction of the velocity v. Thus if the existence of such a vortex were possible, it must hang transverse to v. Contrast footnote, p. 627, which implies an infinite straight vortex: the Newtonian dynamical relativity loses its meaning for infinite systems.

The mode in which the experimentally verified body-vortex completes itself into a ring, whether by shedding a continual trail of vortices from its ends, and perhaps in any case involving its own continual renewal by slip-process, thus opens out a novel and interesting type of pure hydrodynamic problem.]

Digitized by Google

The Pedal (3, 2) *Correspondence.*

By R. VAIDYANATHASWAMY.
(Communicated by Mr H. W. TURNBULL.)

[*Received* 19 November 1926, *read* 31 January 1927.]

CONTENTS.

§ 6. THE PEDAL AND THE SEMI-PEDAL FORMS.
(27) The Semi-Pedal Pencil and the Pedal Form.
(28) The Pedal and the Auto-Form.
(29) The Pedal and the Semi-Pedal Envelopes.
§ 7. THE PEDO-SIMILAR FORMS.
(30) Double specification by measure-deviation.
(31) The co-basal Family of Pedo-Similar Forms.
(32) The net of Pedo-Similar Forms with a given basal triangle.
(33) The structure of the manifold of Pedo-Similar Forms.
§ 8. EXTENSION TO n DIMENSIONS.
(34) The Pedal Correspondence on the Wallace curve.

INTRODUCTION.

This is a sequel to recent communications* to the Society, on "The General (m, n) Correspondence" and "The (2, 1) Correspondence." For convenience the reference numbers here adopted will be consecutive with what has preceded, so that the present work opens with § 6 as a direct consequence of previous results in earlier sections.

The Pedal, Semi-pedal and Pedo-similar correspondences studied in §§ 6, 7 have an individuality of their own, arising from their close connexion with the circular points at infinity. We may contrast the binary field of the circle with the projective properties of conics. The conic of projective geometry has no landmarks; its symmetry is too perfect. But the difference which the entry of the circular points makes in the algebra may be compared with the effect of two powerful centres of disturbance in a uniform field of force. The resulting algebraic processes tend to subordinate the usual symmetries to which we are accustomed, and to reveal new and unforeseen symmetries in their place. In particular, the apolarity conditions illustrate this feature, conditions which naturally connect themselves with the orthocentre of an inscribed triangle. It becomes more and more apparent in the course of the investigation that *the three-dimensional continuum of inscribed triangles of a circle is itself broken up by the circular points into a* $(2+1)$*-dimensional continuum, composed of pedo-parallel systems and orthocentric pencils.*

* *Proc. Camb. Phil. Soc.* XXIII (1926), pp. 109—119, 233—261.

Digitized by Google

§ 6. The Pedal and the Semi-Pedal Forms.

(27) *The Semi-Pedal Pencil and the Pedal Form.*

If $\alpha\beta\gamma$ is an inscribed triangle, the feet of the perpendiculars on its sides from any point x on the circle are collinear; if the line of collinearity cuts the circle in y, the correspondence (x, y) is determined by a (3, 2) form $\Pi(x, y)$, which we call *the pedal form* of $\alpha\beta\gamma$.

We shall prove the collinearity property by means of the idea of Derivate Forms, introduced in (5). It was shewn in 5·17 that if through each p_1 we draw the chord p_1p parallel to the pedal line of p_1 in regard to $\alpha\beta\gamma$, the correspondence (p_1, p) is the (2, 1) polar correspondence $\{f''(x)\}_y$, where $f''(x)$ is given by 5·15. Hence the (3, 2) correspondence (p_1, p_1p) is represented by the defective polar form $(xy)\{f''(x)\}_y$ (1·9). Thus the two values of y corresponding to any x in the (3, 2) correspondence $(xy)\{f''(x)\}_y$ are the extremities of a chord parallel to the pedal line of x. Hence the pedal form $\Pi(x, y)$, considered as a quadratic in y, can only differ by a multiple of $(iy)(jy)$ from $(xy)\{f''(x)\}_y$. Combining with this the fact (which we may be supposed to know from elementary geometry) that α, β, γ are finite fixed points of the pedal correspondence, it is clearly suggested that we should look for the pedal form $\Pi(x, y)$ among the members of the pencil:

6·1

$$\Pi'(x, y) = \lambda(ij)^2(\alpha x)(\beta x)(\gamma x)(iy)(jy) - 4\mu(xy)\{f''(x)\}_y.$$

We shall indeed find that $\Pi'(x, y)$ becomes the pedal form $\Pi(x, y)$ for the values $\lambda = 1$, $\mu = 1$ of the parameters. Before we do this, it is well to notice some obvious properties of the pencil of forms Π' which we call *the semi-pedal forms* of $\alpha\beta\gamma$.

The semi-pedal pencil Π' belongs to the special type of pencil which contains a defective member; hence every semi-pedal form Π' has the same fixed points $\alpha\beta\gamma ij$. From the definition of the derivate form, we see also that the derivate of any semi-pedal form in respect of the partition $\alpha\beta\gamma/ij$, is a multiple of the polar form $\{f''(x)\}_y$. Of particular interest is the behaviour of a semi-pedal correspondence Π' at i, j. We see by actual substitution that $y = i, j$ correspond in any Π' to $x = (i, j, j), (j, j, i)$ respectively. Hence

6·2 (ij) is a closed and complete pair of any semi-pedal correspondence Π' (including by anticipation, the pedal correspondence Π also); or briefly, *any* Π' *is complete and closed at infinity.*

Our method of further procedure is the investigation of the forms which are the derivates of the semi-pedal forms in respect of a partition of the type $\alpha ij/\beta\gamma$. Denoting by Π_α' the derivate of

Digitized by Google

Π' in regard to this partition, we see that Π_α' is the sum of two terms which are the respective derivates of the two terms in the expression 6·1 for $\Pi'(x, y)$. Now the derivate of $(xy)\{f''(x)\}_y$ is simply $\{f''(x)\}_y$; and the derivate of $(\alpha x)(\beta x)(\gamma x)(iy)(jy)$, being the derivate of a form of rank one in respect of a partition not its own, considered in 1·17, is the product of (αx) and the symmetric bilinear form $(\gamma x)(jy)(\beta i) + (\beta y)(ix)(\gamma j)$, which corresponds to the involution determined by the pairs $\beta\gamma$, ij. Thus

6·3

$$\Pi_\alpha'(x, y) = \lambda(ij)^2(\alpha x)\{(\gamma x)(jy)(\beta i) + (\beta y)(ix)(\gamma j)\} - 4\mu\{f''(x)\}_y.$$

Consider the first term in this expression. It is a (2, 1) correspondence in which i, j correspond to each other, that is to say, an auto-correspondence; its fixed points consist of α, and the fixed points of the involution $(\beta\gamma, ij)$, that is, the extremities of the diameter $\alpha_1\alpha_2$ perpendicular to $\beta\gamma$. Thus the first term in 6·3 is the auto-form of the right-angled triangle $\alpha\alpha_1\alpha_2$; the second term is the polar form, and therefore the auto-form, of the equilateral triangle $f''(x)$ (5·2). Thus the forms Π_α' for different values of λ, μ are the auto-forms of triangles of the pencil $\{(\alpha x)(\alpha_1 x)(\alpha_2 x), f''(x)\}$. Now by 5·5 the right-angled triangle $\alpha\alpha_1\alpha_2$ is pedo-perpendicular to $\alpha\beta\gamma$, and therefore pedo-parallel to the triangle $f''(x)$ which is the pedo-perpendicular equilateral triangle of $\alpha\beta\gamma$. Thus the pencil $\{(\alpha x)(\alpha_1 x)(\alpha_2 x), f''(x)\}$ is a pedo-parallel pencil containing an equilateral member $f''(x)$, and therefore a *null* pedo-parallel pencil. Hence

6·4 the pencil Π_α' (which is the derivate of the semi-pedal pencil Π' in regard to the partition $\alpha ij/\beta\gamma$) is the pencil of auto-forms of triangles of the null pedo-parallel pencil, whose orthocentric locus is the diameter through α.

We recall now the four remarkable members of a null pedo-parallel pencil, spoken of in 5·26. Two of the remarkable members in the present case evidently correspond to the two terms in the expression 6·3 for Π_α'. A third remarkable member is the other right-angled triangle in the null pedo-parallel pencil of 6·4, which by 5·26 must have its right angle at the diametrically opposite point of α, and the diameter parallel to $\beta\gamma$ for its hypotenuse. Now the diametrically opposite point of α corresponds to the linear form $(ix)(j\alpha) + (jx)(i\alpha)$, and the involution of chords perpendicular to $\beta\gamma$ to the symmetric bilinear form

$$(ix)(iy)(j\beta)(j\gamma) + (jx)(jy)(i\beta)(i\gamma).$$

Hence the auto-form of this second right-angled triangle is of the shape

$$\{(ix)(j\alpha) + (jx)(i\alpha)\}\{(ix)(iy)(j\beta)(j\gamma) + (jx)(jy)(i\beta)(i\gamma)\}.$$

Digitized by Google

Accordingly, by 6·4, there must exist values of λ, μ which make $\Pi_\alpha'(x, y)$ equal to this auto-form. Using the value 5·15 for $f'''(x)$, we can verify that these values are $\lambda = 1$, $\mu = 1$*. Supposing that $\Pi'(x, y)$ becomes $\Pi(x, y)$ for $\lambda = \mu = 1$, we see then by symmetry that the derivates Π_α, Π_β, Π_γ of $\Pi(x, y)$, in regard to the respective partitions $\alpha ij/\beta\gamma$, $\beta ij/\gamma\alpha$, $\gamma ij/\alpha\beta$, are given by:

6·5

$$\begin{cases} \Pi_\alpha(x, y) = \{(ix)(j\alpha) + (jx)(i\alpha)\}\{(ix)(iy)(j\beta)(j\gamma) \\ \qquad\qquad + (jx)(jy)(i\beta)(i\gamma)\}, \\ \Pi_\beta(x, y) = \{(ix)(j\beta) + (jx)(i\beta)\}\{(ix)(iy)(j\gamma)(j\alpha) \\ \qquad\qquad + (jx)(jy)(i\gamma)(i\alpha)\}, \\ \Pi_\gamma(x, y) = \text{similar expression.} \end{cases}$$

To shew that $\Pi(x, y)$ is the pedal form, we have to shew that it arises from the collinearity of the feet of the perpendiculars. Let x be a given point, y', y'', y''' the values of y which correspond to x in Π_α, Π_β, Π_γ respectively; also let y_1y_2 correspond to x in $\Pi(x, y)$. Since Π_α is the derivate of Π we have a relation of the form

$$\Pi(x, y) = (ij)^2(\alpha x)(ix)(jx)(\beta y)(\gamma y) + (xy)\,\Pi_\alpha(x, y).$$

This relation, interpreted as a linear relation between three binary quadratics in y, clearly implies that the three chords y_1y_2, $\beta\gamma$, xy' are concurrent. Similarly $(y_1y_2, \gamma\alpha, xy'')$, $(y_1y_2, \alpha\beta, xy''')$ are concurrent. But, as has been seen already, xy', xy'', xy''' are perpendicular to $\beta\gamma$, $\gamma\alpha$, $\alpha\beta$ respectively. Therefore the feet of the perpendiculars from x on the sides of $\alpha\beta\gamma$ are collinear, the line of collinearity being y_1y_2, so that $\Pi(x, y)$ is indeed the pedal form.

The different shapes obtained for the pedal form may be set down here for future reference:

6·6

$$\Pi(x, y) = (ij)^2(\alpha x)(\beta x)(\gamma x)(iy)(jy) - 4(xy)\{f''(x)\}_y.$$

* For the verification we may begin by using the following lemma:

If $L_1(x)$, $L_2(x)$ be linear, and $P_1(x)$, $P_2(x)$ quadratic forms, the necessary and sufficient condition that

$$L_1(x)[P_1(x)]_y + L_2(x)[P_2(x)]_y$$

should be a polar form, is

$$(L_1, P_1)' + (L_2, P_2)' \equiv 0.$$

This follows at once by using the development in ladder-shape 1·12, 1·13. By the application of the lemma it follows easily that

$$\{(ix)(j\alpha) + (jx)(i\alpha)\}\{(ix)(iy)(j\beta)(j\gamma) + (jx)(jy)(i\beta)(i\gamma)\} - \lambda(ij)^2(\alpha x)\{(\gamma x)(jy)(\beta i) + (\beta y)(ix)(\gamma j)\}$$

is a polar form only for $\lambda = 1$; it is then clearly the polar form of

$$\{(ix)(j\alpha) + (jx)(i\alpha)\}\{(ix)^2(j\beta)(j\gamma) + (jx)^2(i\beta)(i\gamma)\} - (ij)^2(\alpha x)\{(\gamma x)(jx)(\beta i) + (\beta x)(ix)(\gamma j)\}.$$

This expression may be verified to be equal to $-4f''(x)$, where $f''(x)$ is given by 5·15.

Digitized by Google

6·7

$$\Pi(x, y) = (ij)^2(\alpha x)(ix)(jx)(\beta y)(\gamma y) + (xy)\,\Pi_\alpha(x, y),$$
$$= (ij)^2(\beta x)(ix)(jx)(\gamma y)(\alpha y) + (xy)\,\Pi_\beta(x, y),$$
$$= (ij)^2(\gamma x)(ix)(jx)(\alpha y)(\beta y) + (xy)\,\Pi_\gamma(x, y),$$

where $f'''(x)$ is given by 5·15, and Π_α, Π_β, Π_γ by 6·5.

It was seen in 6·2 that the pedal correspondence has a particular type of closure at infinity; namely $y = i, j$ correspond in it to $x = (ijj), (jii)$ respectively. This type of closure does not characterise exclusively the pedal or even the semi-pedal correspondences; it belongs in fact to the family of $\infty^{11-6} = \infty^5$ '*pedo-similar*' correspondences studied in the next section. The exclusive property of the pedal correspondences Π among these latter is that, if αa be the diameter through α, $x = a$ corresponds in Π to $y = \beta, \gamma$.

(28) *The Pedal Form and the Auto-Form.*

From 3·6, the (1, 2) auto-form of $\alpha\beta\gamma$, namely the form which has ij for auto-pair, and $\alpha\beta\gamma$ for its fixed points, is shewn to be of the shape

$$\Sigma(\alpha i)(\alpha j)(\beta\gamma)(\beta\gamma, ij)(\alpha x)(\beta y)(\gamma y),$$

where
$$(\beta\gamma, ij) = (\beta i)(\gamma j) + (\beta j)(\gamma i).$$

We shall find it convenient to denote by $A(x, y)$ a certain multiple of this auto-form; namely

6·8

$$A(x, y) = \frac{2}{(\beta\gamma)(\gamma\alpha)(\alpha\beta)}\,\Sigma(\alpha i)(\alpha j)(\beta\gamma)(\beta\gamma, ij)(\alpha x)(\beta y)(\gamma y).$$

We notice that

$$A(x, x) = 2(ij)^2(\alpha x)(\beta x)(\gamma x).$$

Putting now $x = i$ in 6·7, we have

$$\Pi(i, y) = (ij)^2(i\alpha)(i\beta)(i\gamma)(iy)(jy).$$

Hence
$$\Pi(x, y) + (ij)^2(\alpha x)(\beta x)(\gamma x)(iy)(jy)$$

vanishes when $x = i$ or j. Therefore there is an identity of the form

6·9

$$\Pi(x, y) + (ij)^2(\alpha x)(\beta x)(\gamma x)(iy)(jy) \equiv \kappa(ix)(jx)\,A(x, y),$$

where κ is numerical, and, from the closure property of the pedal form, $A(x, y)$ is simply the auto-form 6·8 of $\alpha\beta\gamma$. Putting $y = x$, we see that κ must be equal to 1.

This relation 6·9 may be reached in a more instructive way. It was seen in 6·4 that the derivate pencil Π_a' of the semi-pedal pencil Π' consisted of the auto-forms of members of a null pedo-

Digitized by Google

parallel pencil, and *three* semi-pedal forms (namely, the pedal form and the two terms in the expression 6·1) were found which corresponded to three of the four remarkable members of the null pedo-parallel pencil. The fourth remarkable member is that of the shape $(ix)(jx)(\lambda x)$, the auto-form of which is $(ix)(jx)(\lambda y)$. Now $\Pi'(x, y)$ can have the derivate $\Pi_x' = (ix)(jx)(\lambda y)$ only if Π' has the factor $(ix)(jx)$. Thus there is a semi-pedal form $(ix)(jx)A(x, y)$, having the factor $(ix)(jx)$; and from the closure property of Π', $A(x, y)$ must be an auto-form. The harmonic property of the four remarkable members of the null pedo-parallel pencil (5·26) easily shews that the values of λ, μ in 6·1 which give this semi-pedal form are $\lambda : \mu = 2 : 1$. We may now verify that

$$6\cdot91 \qquad \begin{aligned} &2(ij)^2(\alpha x)(\beta x)(\gamma x)(iy)(jy) - 4(xy)\{f''(x)\}_y \\ &= \Pi(x, y) + (ij)^2(\alpha x)(\beta x)(\gamma x)(iy)(jy) \\ &= (ix)(jx)A(x, y), \end{aligned}$$

where $A(x, y)$ is given by 6·8.

We have in (27), (28) reached a fundamental theorem of our subject, which we may set forth as follows:

6·10 *The semi-pedal pencil of* $\alpha\beta\gamma$ *contains, like the null pedo-parallel pencil, two harmonic pairs of remarkable members; one of these pairs is composed of the pedal form* $\Pi(x, y)$ *and the form of rank one,* $(\alpha x)(\beta x)(\gamma x)(iy)(jy)$, *the other, of the defective polar form of the pedo-perpendicular equilateral triangle of* $\alpha\beta\gamma$, *and a form of rank two, namely the product of* $(ix)(jx)$ *and the* (1, 2) *auto-form of* $(\alpha\beta\gamma)$.

The geometrical meaning of the harmonic property of these four members is obviously, that the pedal line of x bisects the join of x with the orthocentre of $\alpha\beta\gamma$.

(29) *The Pedal and the Semi-Pedal Envelopes.*

From (8) it follows that the pedal envelope (namely, the envelope of the pedal lines of $\alpha\beta\gamma$) is a rational envelope of the third class, with the line at infinity for its bitangent. From the closure property, the points of contact of the bitangent are i, j: hence the envelope is a three-cusped hypocycloid. It may also be seen that the pedal line of a point near i is the tangent to the envelope at a point near j. *All this is equally true of any semi-pedal correspondence.*

The concurrency system of the pedal correspondence of $\alpha\beta\gamma$ is the system of triads of points, whose pedal lines are concurrent. It is clearly a linear ∞^2-system. The pedal lines of i, j being each the line at infinity, it follows that the concurrency system contains the singular pencil $ij\lambda$, and is therefore a pedo-parallel system. Now the chords which correspond to α, β, γ in the pedal, or in any

Digitized by Google

semi-pedal correspondence of $\alpha\beta\gamma$, are simply the perpendiculars of the triangle $\alpha\beta\gamma$, which are concurrent at the orthocentre. Thus $\alpha\beta\gamma$ belongs to the concurrency system of every semi-pedal form defined by it. We have therefore a theorem which enlarges still further the content of "pedo-parallelism":

6·11 *the concurrency system of the pedal or any semi-pedal form of* $\alpha\beta\gamma$ *is the pedo-parallel system of* $\alpha\beta\gamma$*.

From 6·1, 6·9 a semi-pedal form $\Pi'(x, y)$ can be written

6·12

$$\Pi'(x, y) = \lambda(ix)(jx)A(x, y) - \kappa\lambda(ij)^2(\alpha x)(\beta x)(\gamma x)(iy)(jy).$$

We shall term κ the *measure* of Π'; we shall also call the chord y_1y_2 which corresponds to any x in Π', the *κ-pedal line* of x in respect of $\alpha\beta\gamma$, and its envelope the *κ-pedal envelope*. By comparing 6·9, 6·12, it follows that the κ-pedal line of x is the line parallel to, and κ times as far from the orthocentre as, the pedal line of x. Hence

6·13 *the κ-pedal envelope results on uniform dilatation of the pedal envelope from the orthocentre, the constant of dilatation being κ.*

In particular, for $\kappa = \infty$ we have the form $(\alpha x)(\beta x)(\gamma x)(iy)(jy)$, whose envelope lies wholly at infinity. For $\kappa = 0$ we have the form $(ix)(jx)A(x, y)$, whose envelope is the orthocentre and the circular points. For $\kappa = 2$ we have the envelope E of the defective polar form, defined by the pedo-perpendicular equilateral triangle $\alpha''\beta''\gamma''$, of $\alpha\beta\gamma$.

The three tangents to E from any point x on the circle are the chords joining x to its second polar point x' and to its first polar points x_1x_2, in respect of $\alpha''\beta''\gamma''$. Hence *one* of the three tangents to E from any point on the circle is algebraically distinct from the other two. It can be shewn from general reasoning that this necessitates that E should have triple contact with the circle. As a matter of fact we see directly that it has triple contact at $\alpha''\beta''\gamma''$. From 6·11 it follows that the cuspidal tangents of E are the diameters $\alpha'\alpha''$, $\beta'\beta''$, $\gamma'\gamma''$, where $\alpha'\beta'\gamma'$ is the pedo-parallel equilateral triangle of $\alpha\beta\gamma$.

6·14 The envelope E is also the locus of orthocentres of line-triangles pedo-parallel to $\alpha\beta\gamma$.

* Thus the concurrency-systems of all semi-pedal forms of $\alpha\beta\gamma$ are the same. 5·18 shews the sameness of the concurrency-systems for two particular semi-pedal forms, namely the pedal form and the defective polar form. The sameness of the concurrency-systems is also a direct consequence of 6·13.

It follows from 1·1 and 6·11 that the second rank covariant of any semi-pedal form of $\alpha\beta\gamma$ is a binary cubic whose roots correspond to the vertices of the pedo-perpendicular equilateral triangle of $\alpha\beta\gamma$.

Digitized by Google

For, let X_1X_2 be a diameter, and let XXX_1 be a line-triangle pedo-parallel to $\alpha\beta\gamma$. The pedal line of X is then perpendicular to XX_1 (5·8), and therefore parallel to XX_2. Hence by 5·17 XX_2 touches E; from 5·1 it may be shewn that the point of contact is the reflection of X_2 in X. But the reflection of X_2 in X is the orthocentre of the line triangle XXX_1. Hence the theorem.

As an immediate application, the cusps of E being the orthocentres of the point-triangles $\alpha'\alpha'\alpha'$, $\beta'\beta'\beta'$, $\gamma'\gamma'\gamma'$ lie on $S\alpha'$, $S\beta'$, $S\gamma'$, at distances 3ρ from S (5·13). By 6·13 the pedal envelope is obtained on contracting E from O, to half its linear dimensions. Hence:

6·15 *The nine-point circle of $\alpha\beta\gamma$ is the in-circle of its pedal envelope, the points of contact being the midpoints of $O\alpha''$, $O\beta''$, $O\gamma''$.* The cusps of the pedal envelope are situated on radii vectores from the nine-point centre N parallel to $S\alpha'$, $S\beta'$, $S\gamma'$, at a distance $\frac{3}{2}\rho$ from it. The pedal envelope when shifted through a distance $\frac{1}{2}OS$ parallel to OS becomes the locus of nine-point centres of line-triangles pedo-parallel to $\alpha\beta\gamma$. The point of contact of the pedal line of x with its envelope lies therefore on the join of O to the orthocentre of the line-triangle xxx' pedo-parallel to $\alpha\beta\gamma$.

The algebraic distinctness of one of the tangents to E' from any point on the circle from the others reappears after contraction as the distinctness of one of the three pedal lines through any point on the nine-point circle, from the other two; the distinctness arises from the proposition that *the pedal lines of the extremities of a diameter meet at right angles on the nine-point circle*, the third pedal line through their intersection being consequently of an algebraically different character from them. From 6·13, 6·15 we now easily obtain:

6·16 *The centre of the κ-pedal envelope is in OS at a distance $\frac{1}{2}\kappa\, OS$ from O.* If the pedal angle between two co-orthocentric triangles is θ, the κ-pedal envelope of one is obtained by turning that of the other through an angle $2\theta/3$ about its centre. Thus for a given κ, the κ-pedal envelope has always the same dimensions, but takes its ∞^3 different positions in the plane for the ∞^3 inscribed triangles $\alpha\beta\gamma$. The position of the centre of the envelope indicates the particular orthocentric pencil to which $\alpha\beta\gamma$ belongs, while the disposition of the envelope about its centre indicates the pedo-parallel system to which $\alpha\beta\gamma$ belongs.

§ 7. The Pedo-Similar Forms.

(30) *Double specification of measure-deviation.*

The most general (3, 2) forms which have closure at infinity in the manner of the pedal form (that is, in which $y = i.\, j$ correspond

Digitized by Google

to $x = ijj, jii$ respectively) we call "the pedo-similar forms." The pedo-similar forms involve $12 - 6 = 6$ parameters, and include as special cases the semi-pedal and the pedal forms. Thus there are ∞^3 pedal correspondences, ∞^4 semi-pedal correspondences, and ∞^5 pedo-similar correspondences. Since the angle-property 5·1 is simply a consequence of the behaviour at i, j, we see that

7·1 *The angle-property* 5·1 *holds for the general pedo-similar form.*

This theorem suggests a specification of any pedo-similar form $\Pi''(x, y)$ by means of a *basal triangle* $\alpha\beta\gamma$ (which, be it noted, is not in general the fixed triangle of Π''), a parent semi-pedal form $\Pi'(x, y)$ of measure κ, defined by $\alpha\beta\gamma$, and a "deviation" θ. Let x be any point on the circle, x' the point such that $x\hat{S}x'$ is equal to the angle of deviation θ. If y_1y_2 be the κ-pedal line of x' in respect of $\alpha\beta\gamma$, then the correspondence (x, y_1y_2) is clearly a pedo-similar correspondence; it is defined in respect of the basal triangle $\alpha\beta\gamma$, and is said to be of measure-deviation (κ, θ). We shall also call y_1y_2 the (κ, θ)-pedal line of x in regard to $\alpha\beta\gamma$. Since there are two parameters κ, θ in addition to the three in the choice of the basal triangle, it follows that all pedo-similar correspondences will be covered by this specification. Also, from our construction, the set of (κ, θ)-pedal lines of $\alpha\beta\gamma$ is composed simply of the κ-pedal lines of $\alpha\beta\gamma$. Hence

7·2 *The envelope associated with the pedo-similar form, with the basal triangle* $\alpha\beta\gamma$, *and the measure-deviation* (κ, θ), *is simply the* κ*-pedal envelope of* $\alpha\beta\gamma$.

Hence, as in 6·11, the concurrency system of the pedo-similar form is a pedo-parallel system; if we now rotate $\alpha\beta\gamma$ through an angle $-\theta$ about the centre, into the position $\alpha'\beta'\gamma'$, then the (κ, θ)-pedal lines of α', β', γ' are clearly the κ-pedal lines of α, β, γ, which, being the perpendiculars of the triangle $\alpha\beta\gamma$, are concurrent. Thus $\alpha'\beta'\gamma'$ belongs to the concurrency system. Since the pedal angle between $\alpha\beta\gamma, \alpha'\beta'\gamma'$ is $-\frac{3}{2}\theta$,

7·3 *The concurrency system of a pedo-similar form of deviation* θ *is the pedo-parallel system which makes the pedal angle* $-\frac{3}{2}\theta$ *with its basal triangle.*

We have now to examine in how many ways a given pedo-similar correspondence can be specified in this manner by a basal triangle Δ, a measure κ, and a deviation θ. We shall find that there are precisely two specifications $(\Delta_1, \kappa_1, \theta_1), (\Delta_2, \kappa_2, \theta_2)$, where $\kappa_2 = -\kappa_1, \theta_2 = \theta_1 + \pi$; the two possible basal triangles Δ_1, Δ_2 we shall call *associated triangles.*

Digitized by Google

The theory of associated triangles depends on the following important property of the null pedo-parallel pencil:

7·4 *If $\alpha\beta\gamma$ is an inscribed triangle, there are precisely three points a, b, c which have the property that their κ-pedal lines in respect of $\alpha\beta\gamma$ pass through their diametrically opposite points a', b', c'.* For different values of κ, the triangle abc belongs to the null pedo-parallel pencil which contains $\alpha\beta\gamma$. Conversely, any two members $\alpha\beta\gamma$, abc of a null pedo-parallel pencil are such that there is a number κ, such that the κ-pedal lines of a, b, c in regard to $\alpha\beta\gamma$ pass through their diametrically opposite points.

That there are exactly three points a, b, c is clear from the fact that a', b', c' are the finite fixed points of the (κ, π) pedo-similar correspondence with the basal triangle $\alpha\beta\gamma$. To prove that the pencil $\{\alpha\beta\gamma, abc\}$ is a null pedo-parallel pencil, let O be the orthocentre of $\alpha\beta\gamma$, and let our fundamental circle be called C. Let x be any point on C, and let Ox be divided at ξ in the ratio $\kappa : 2-\kappa$. The locus of ξ is then a circle C'; O is clearly one of the two centres of similitude of C, C', let O_1 be the other, so that if S' be the centre of C', O and O_1 divide $S'S$ in the ratios $-\frac{1}{2}\kappa, \frac{1}{2}\kappa$. From the property of the centre of similitude it then follows that the join of ξ to the diametrically opposite point of x passes through O_1 for all positions of x. But the join of ξ to the diametrically opposite point of x is the κ-pedal line of x, when x is at a, b or c. Thus the κ-pedal lines of a, b, c are concurrent at O_1, and therefore by 6·11 abc is pedo-parallel to $\alpha\beta\gamma$.

It was seen in (25) that the members of the pedo-parallel system of $\alpha\beta\gamma$ are in linear correspondence $\mathfrak{A}$ with their orthocentres. It is easy to see that these members are also in linear correspondence $\mathfrak{A}_\kappa$ with the point of concurrence of the κ-pedal lines of their vertices in regard to $\alpha\beta\gamma$. It is immediately verified that any member $ij\lambda$ corresponds to the point at infinity on the pedal line of λ in $\mathfrak{A}$ as well as in $\mathfrak{A}_\kappa$. From this it easily follows that the point of concurrence of the κ-pedal lines of abc is obtained from the orthocentre of abc by a uniform radial dilatation from the orthocentre of $\alpha\beta\gamma$. Thus we have the theorem:

7·5 *The join of the orthocentres of two pedo-parallel triangles $\alpha\beta\gamma$, abc contains the points of concurrence of the κ-pedal lines of the vertices of either in regard to the other.* Hence the orthocentre of abc lies in the line OO_1, that is, in the diameter through O. The orthocentric locus of the pencil $\{\alpha\beta\gamma, abc\}$, which must be a line (5·24), is thus shewn to be a *diameter*, so that the pencil $\{\alpha\beta\gamma, abc\}$ is a null pedo-parallel pencil (cf. (25)). We shall say that the triangle $a'b'c'$ (where a', b', c' are the diametrically opposite points of a, b, c) is the triangle *associated with $\alpha\beta\gamma$ in respect of the*

Digitized by Google

modulus κ. Now the pedal angle, as well as the anti-pedal angle, between two diametrically opposite triangles abc, $a'b'c'$ is $\frac{1}{2}\pi$; that is, abc and $a'b'c'$ are pedo-perpendicular, as well as anti-pedo-perpendicular. But, since $\alpha\beta\gamma$, abc belong to a null pedo-parallel pencil, their pedal as well as their anti-pedal angle vanishes. Hence

7·6 The triangles associated, modulus κ, with $\alpha\beta\gamma$ are, for different values of κ, the triangles diametrically opposite to the members of the null pedo-parallel pencil containing $\alpha\beta\gamma$; thus associated triangles are pedo-perpendicular, as well as anti-pedo-perpendicular.

This indicates that the relation between associated triangles is a symmetrical one. More precisely, we have the result:

7·7 *If $a'b'c'$ is associated, modulus κ, with $\alpha\beta\gamma$, then $\alpha\beta\gamma$ is associated, modulus $-\kappa$, with $a'b'c'$.*

For, it is clear from the symmetry of 7·6 that $\alpha\beta\gamma$ must be associated in respect of some modulus κ' with $a'b'c'$. Here κ' must be a rational function $R(\kappa)$ of κ, such that $R\{R(\kappa)\}=\kappa$ identically. Hence R must be a homographic involution. Now the O-pedal line is the auto-chord, and therefore the triangle associated, modulus zero, with $\alpha\beta\gamma$ is the co-orthocentric pedo-perpendicular triangle of $\alpha\beta\gamma$. Again the 2-pedal line of x in regard to $\alpha\beta\gamma$ is the join of x to its second polar point in respect of its pedo-perpendicular equilateral triangle $\alpha''\beta''\gamma''$ (5·17, 6·10). It follows that the triangle associated, modulus 2, with $\alpha\beta\gamma$ is $\alpha''\beta''\gamma''$. Now the pedal line of α in regard to $\alpha''\beta''\gamma''$ bisects $S\alpha$, and is parallel to $\beta\gamma$ (5·8). Hence the κ'-pedal line of α will pass through the diametrically opposite point, if $\kappa'=-2$. Thus $R(0)=0$, and $R(2)=-2$. Hence $R(\kappa)\equiv-\kappa$.

We may now state the fundamental property of associated triangles:

7·8 If x, x' be variable points on the circle, such that $x\hat{S}x'=\theta$, the κ-pedal line of x in regard to Δ_1 can be, for all positions of x, the κ'-pedal line of x' in regard to a second inscribed triangle Δ_2, only if $\theta=\pi$; then κ' must be equal to $-\kappa$, and Δ_2 is the triangle associated, modulus κ, with Δ_1.

We easily see that there are only three positions x_1, x_2, x_3 of x whose κ-pedal lines pass through the corresponding positions x_1', x_2', x_3' of x'. Hence if a second triangle Δ_2 with the required property exists at all, it must be $x_1'x_2'x_3'$. Now the κ'-pedal lines of x_1', x_2', x_3' in regard to $x_1'x_2'x_3'$ are the perpendiculars of the triangle $x_1'x_2'x_3'$, and are concurrent; hence the κ-pedals of x_1, x_2, x_3 in regard to Δ_1 are concurrent, so that $x_1x_2x_3$ is pedo-parallel to

Digitized by Google

Δ_1 (6·11). This involves that the κ-pedal line of x_1 should be perpendicular to x_2x_3 (5·8), so that the triangles $x_1x_2x_3$, $x_1'x_2'x_3'$ have their corresponding sides parallel. This cannot be unless $\theta = \pi$.

If $\theta = \pi$, then from our definition $\Delta_2 = x_1'x_2'x_3'$ is the triangle associated, modulus κ, with Δ_1, and by 7·7, Δ_1 is associated, modulus $-\kappa$, with Δ_2.

Now supposing that x' is the diametrically opposite point of x, if $(x \to y)$ is the κ-pedal correspondence of Δ_1, then the correspondence $(x' \to y)$ must be a semi-pedal correspondence of Δ_2 (since the chords corresponding to $x_1'x_2'x_3'$ are perpendicular to their opposite sides). Suppose then, that it is the κ'-pedal correspondence of Δ_2. From the reciprocity of the triangles Δ_1, Δ_2 we can shew as in 7·7 that $\kappa' = -\kappa$. We may also shew the same thing directly by considering the κ-pedal line of either extremity of the diameter containing the orthocentres of Δ_1, Δ_2. Thus 7·8 is proved; it also follows that

7·9 The (κ, θ_1)-pedo-similar correspondence of the basal triangle Δ_1 can be expressed in only one other way as the (κ', θ_2)-pedo-similar correspondence of Δ_2; $\kappa' = -\kappa$, $\theta_2 = \theta_1 + \pi$, and Δ_2 is the triangle associated, modulus κ, with Δ_1.

It is seen from this that the deviation θ of a given pedo-similar correspondence is determinate to the modulus 2π, when the corresponding basal triangle is given, but determinate only to the modulus π when it is not given.

(31) *The co-basal Family of Pedo-Similar Forms.*

If x, x' are points on the circle such that $x\hat{S}x' = \theta$, then since $(ij)x = (ix)j - (jx)i$, we can take

$$x' = e^{\sqrt{-1}\frac{\theta}{2}}(ix)j - e^{-\sqrt{-1}\frac{\theta}{2}}(jx)i = (ij)x + \sqrt{-1}\tan\tfrac{1}{2}\theta[(ix)j + (jx)i],$$

discarding the scalar factor $\cos\frac{1}{2}\theta$. Hence the (κ, θ)-pedo-similar form defined by the basal triangle $\alpha\beta\gamma$ can be taken to be

$$7\cdot10 \qquad \Pi''(x, y) = \Pi'(x', y)$$
$$= \Pi(x', y) - (\kappa - 1)(ij)^2(\alpha x')(\beta x')(\gamma x')(iy)(jy),$$

where $$x' = (ij)x + \sqrt{-1}\tan\tfrac{1}{2}\theta\{(ix)j + (jx)i\}.$$

We notice that Π'' involves the parameter κ linearly, and the parameter $\tan\frac{1}{2}\theta$ to the third degree. Hence

7·11 The fixed triangles of co-basal pedo-similar forms Π'', of given deviation θ, belong to a pencil F_θ; the first triangles of

Digitized by Google

co-basal pedo-similar forms of given measure κ constitute a cubic family G_κ. Since the semi-pedal form of measure κ has $\alpha\beta\gamma$ for its fixed triangle, it follows that G_κ contains the basal triangle for all values of κ.

The system G_∞ is the set of fixed triangles of forms of the shape $(ax')\,(\beta x')\,(\gamma x')(iy)\,(jy)$, that is, it is the congruent system of $\alpha\beta\gamma$ studied in (26). There are two values of κ for which G_κ apparently becomes a pencil. It is clear first of all, that the pedo-similar forms of zero measure are, from the definition, simply the products of (ix), (jx) with auto-forms of triangles co-orthocentric with $\alpha\beta\gamma$. To be more precise, we see from the angle-property 5·1 that the fixed triangle of the $(0, \theta)$-pedo-similar form is that co-orthocentric triangle which makes the pedal angle $-\frac{1}{2}\theta$ with the basal triangle. Thus:

7·12 G_0 is apparently the orthocentric pencil of the basal triangle $\alpha\beta\gamma$. The fixed triangle of the $(0, \theta)$-pedo-similar form is that co-orthocentric triangle which makes the pedal angle $-\frac{1}{2}\theta$ with $\alpha\beta\gamma$.

Again the semi-pedal form of measure 2 is the defective polar form; consider the pedo-similar form of measure 2 and deviation θ. The fixed points of the form are clearly the points ϵ such that the pedal of ϵ' is parallel to $\epsilon\epsilon'$, where $\epsilon\hat{S}\epsilon' = \theta$, from which it may be easily shewn that the fixed triangle is an equilateral triangle which makes the pedal angle $\frac{1}{2}\pi - \theta$ with $\alpha\beta\gamma$. Hence:

7·13 The system G_2 is apparently the pencil of equilateral triangles; the fixed triangle of the $(2, \theta)$-pedo-similar form is that equilateral triangle which makes the pedal angle $\frac{1}{2}\pi - \theta$ with the basal triangle.

Any pencil F_θ has a member in common with each G_κ; in particular, from 7·12, 7·13, we see that F_θ is the pencil determined by the two triangles $\Delta_{-\frac{1}{2}\theta}$, $E_{\frac{1}{2}\pi-\theta}$, which are respectively the co-orthocentric and equilateral triangles making the pedal angles indicated in the suffix, with $\alpha\beta\gamma$. Now since F_θ contains an equilateral triangle, its members make a constant anti-pedal angle with the basal triangle (26). By 5·29, we see that this constant anti-pedal angle is simply the pedal angle between the co-orthocentric triangles $\Delta_{-\frac{1}{2}\theta}$ and $\alpha\beta\gamma$. Hence we have the theorem:

7·14 *The anti-pedal angle between the basal triangle, and the fixed triangle of a pedo-similar form of deviation* θ, *is* $-\frac{1}{2}\theta$. To obtain the orthocentric locus of F_θ, let O be the orthocentre of $\alpha\beta\gamma$ and O' the orthocentre of the triangle $\alpha'\beta'\gamma'$ obtained by turning

Digitized by Google

$\alpha\beta\gamma$ through an angle $-\theta$ about S. Then $SO = SO'$, and $O\hat{S}O' = -\theta$. By 7·12, 7·13, F_θ contains an equilateral triangle and a co-orthocentric triangle of $\alpha\beta\gamma$, so that its orthocentric locus is a circle (5·24), passing through S, O. Also the (∞, θ)-pedo-similar form being $(\alpha x')\,(\beta x')\,(\gamma x')\,(iy)\,(jy)$ its fixed triangle is $\alpha'\beta'\gamma'$. Therefore

7·15 The orthocentric locus of the pencil F_θ is the circle SOO'. The pedal angle is determinate to the modulus π, while, since the basal triangle is known, the deviation is determinate to the modulus 2π. Hence, if a member of G_0 is given as the fixed triangle, then θ is determined modulo 2π by 7·12, and therefore the pedo-similar form and the corresponding pencil F_θ are determined uniquely. But if a member of G_2 is given as the fixed triangle, then 7·13 determines θ only to the modulus π, and therefore there are two pencils F_θ of the form F_{θ_1}, $F_{\theta_1+\pi}$. Hence

7·16 In general the pencils F_θ establish a (1, 1) correspondence between any two systems G_κ, in particular between G_0 and G_κ. The case $\kappa = 2$ is however exceptional; the pencils F_θ establish a (1, 2) correspondence between G_2 and any G_κ, in particular, between G_2 and G_0. Hence *the totality of fixed triangles of a co-basal family of pedo-similar forms constitute a ruled-cubic system.*

These results can all be verified directly by a geometrical representation of the type indicated in (26). For this purpose we write the (κ, θ) in the shape 7·10, namely

$$\Pi''(x, y) = \Pi'(x', y)$$
$$= \Pi(x', y) - (\kappa - 1)(ij)^2(\alpha x')(\beta x')(\gamma x')(iy)(jy),$$

where $x' = \lambda(jx)i + \mu(ix)j$, so that $\lambda + e^{\sqrt{-1}\theta}\mu = 0$.

Using 6·7 we find that the fixed cubic $\Pi''(x, x) = \Pi'(x', x)$ reduces to

7·17 $$X_0(ix)^3(\alpha j)(\beta j)(\gamma j) + X_1(ix)^2(jx)\Sigma(\alpha i)(\beta j)(\gamma j)$$
$$+ X_2(ix)(jx)^2\Sigma(\alpha i)(\beta i)(\gamma j) + X_3(jx)^3(\alpha i)(\beta i)(\gamma i),$$

where

7·18 $$\begin{cases} X_0 = -\mu^2(2\lambda + \kappa\mu), \\ X_1 = (2 - \kappa)\lambda\mu^2, \\ X_2 = (2 - \kappa)\lambda^2\mu, \\ X_3 = -\lambda^2(2\mu + \kappa\lambda). \end{cases}$$

Digitized by Google

If we represent the cubic 7·17 by the point $(X_0X_1X_2X_3)$ of three-space, the tetrahedron of reference would clearly be the osculating tetrahedron at I, J to the fundamental cubic C of (26), namely the tetrahedron whose edges are the chord IJ, the axis IJ, and the tangents at I, J to C. The equations 7·18 are then the parametric equations of the cubic scroll:

7·19 $$(X_0 + X_1)\,X_2^2 = (X_2 + X_3)\,X_1^2.$$

Supposing λ, μ to be fixed, the equations 7·18 give the generators F_θ of this scroll. Supposing κ to be fixed, the same equations are the parametric equations of a twisted cubic G_κ. We easily verify that G_μ passes through the point corresponding to the basal triangle, namely the point $(1, -1, 1, -1)$, and that it passes through the points I, J (which are the points X_0, X_3 of reference) and has the same tangents at I, J as the fundamental cubic C (namely X_0X_1, X_3X_2). Thus the cubics G_κ pass through five points, so that any two of them constitute the complete intersection of the cubic scroll with a quadric. One particular G_κ, namely G_∞, represents the congruent system of the basal triangle, and therefore has not only the same tangents, but also the same osculating planes at IJ as C.

For $\kappa = 2$, G_κ becomes the chord IJ (hitherto called G_2 (7·13)), counted twice, and the line F_θ; for $\kappa = 0$, G_κ becomes the line through $(1, -1, 1, -1)$ meeting the tangents X_0X_1, X_3X_2 at IJ (hitherto called G_0) and the tangents at I, J. Thus G_2 (that is, the chord IJ) is the double line of the scroll, and as already proved, the generators F_θ establish a (1, 2) correspondence between G_2 and G_0.

(32) *The net of Pedo-Similar Forms with a given basal triangle.*

Since the pedo-similar correspondences with a given fixed triangle $\alpha\beta\gamma$ constitute a linear ∞^2-system, it follows that the totality of defective pedo-similar forms must constitute a pencil contained in every such system (4). But we know that the defective polar forms of equilateral triangles are semi-pedal forms, and do constitute a pencil. We conclude therefore that the only defective pedo-similar forms are the defective polar forms of equilateral triangles. Since $(\alpha x)\,(\beta x)\,(\gamma x)\,(iy)\,(jy)$ is a particular semi-pedal form with the fixed triangle $\alpha\beta\gamma$, it follows that the general pedo-similar form with the same fixed triangle differs from $(\alpha x)\,(\beta x)\,(\gamma x)\,(iy)\,(jy)$ by the defective polar form of an arbitrary equilateral triangle; otherwise expressed, the derivates in respect of the partition $\alpha\beta\gamma/ij$ of forms of the pedo-similar net determined by the fixed triangle $\alpha\beta\gamma$ are the polar forms of equilateral triangles.

7·20 If $\alpha\alpha'$ be the chord perpendicular to $\beta\gamma$, the derivates in

Digitized by Google

respect of the partition $\alpha ij/\beta\gamma$ of forms of the same pedo-similar net, are the auto-forms of triangles of the equilateral system determined by the Euler point α'.

This theorem is the natural extension of 6·4; for, the semi-pedal pencil of $\alpha\beta\gamma$ is obviously contained in the pedo-similar net determined by the fixed triangle $\alpha\beta\gamma$, and we can easily verify that the null pedo-parallel pencil of 6·4 is contained in the equilateral system of 7·20. To prove 7·20, we see from 1·17 *et seq.* that the derivates in respect of $\alpha ij/\beta\gamma$ differ from the derivates in respect of $\alpha\beta\gamma/ij$, namely the polar forms of equilateral triangles by a multiple of the derivate of $(\alpha x)(\beta x)(\gamma x)(iy)(jy)$ in respect of $\alpha ij/\beta\gamma$. This last derivate was seen in (27) to be the auto-form of a right-angled triangle, whose right angle is at α and whose hypotenuse is the diameter perpendicular to $\beta\gamma$. Hence the derivates in respect of $\alpha ij/\beta\gamma$ of forms of the net are auto-forms of triangles of the equilateral system containing this right-angled triangle. Since the Euler point of a right-angled triangle is the extremity of the chord through the right angle parallel to the hypotenuse, we have the result stated, from 5·27.

The problem of finding the basal triangle when the fixed triangle $\alpha\beta\gamma$ and the measure-deviation (κ, θ) are known can be solved analytically by adapting equations 7·18 to a slightly different coordinate system. Let the point $X_0X_1X_2X_3$ now represent the cubic $\Sigma X_r (ix)^{3-r}(jx)^r$; let $X_0X_1X_2X_3$ correspond to the fixed triangle and $A_0A_1A_2A_3$ to the basal triangle. It may be shewn that 7·19 now becomes

7·21 $$(A_1X_0 - A_0X_1)X_2^2A_1A_3 = (A_3X_2 - A_2X_3)X_1^2A_0A_2,$$

and the equations 7·18 can now be written

7·22 $$\begin{cases} A_0 = -\lambda^2 X_0 (2-\kappa)(2\mu+\kappa\lambda), \\ A_1 = -\lambda X_1 (2\lambda+\kappa\mu)(2\mu+\kappa\lambda), \\ A_2 = \mu X_2 (2\lambda+\kappa\mu)(2\mu+\kappa\lambda), \\ A_3 = \mu^2 X_3 (2-\kappa)(2\lambda+\kappa\mu). \end{cases}$$

These determine the basal triangle. When the fixed triangle and the deviation θ are known, 7·22 shews that the locus of $A_0A_1A_2A_3$ is a conic through the points of reference X_0X_3 and through a point in X_1X_2. Thus the system of basal triangles $A_0A_1A_2A_3$ is a conic-system of the type mentioned in 5·24. Hence by 5·24,

7·23 *The orthocentric locus of the basal triangles of pedo-similar forms of deviation* θ *and fixed triangle* $\alpha\beta\gamma$ *is the line* OO', *where* O *is the orthocentre of* $\alpha\beta\gamma$ *and* O' *is such that*

$$SO' = SO, \quad O\hat{S}O' = \theta^*.$$

* The theorem 7·23 may also be obtained by elementary geometry as a necessary consequence of 7·15.

Digitized by Google

We may also find the basal triangle in a more instructive way. Consider first the following problem:

7·24 The κ-pedal line of a given point p in respect of an unknown inscribed triangle $\alpha'\beta'\gamma'$ passes through a given point Q. It is required to specify $\alpha'\beta'\gamma'$.

It is clear that the possible triangles $\alpha'\beta'\gamma'$ belong to a linear ∞^2-system L, and are therefore apolar to a certain triangle $\alpha\beta\gamma$. We specify L by means of its equilateral member and its orthocentric pencil (25). Join pQ and produce it to O, so that

$$OQ : Op = \kappa : 2.$$

The orthocentric pencil of O clearly belongs to L. Also join Sp and divide it at Q', so that $SQ' : Sp = \kappa : 2$. Then the particular equilateral triangle in respect of which the pedal line of p is parallel to QQ', is clearly the required equilateral member of L. Thus L is known. Now the assigning of the fixed triangle abc of a (κ, θ)-pedo-similar form Π'' is equivalent to the statement that the κ-pedal lines of $a'b'c'$ in regard to the basal triangle of Π'' pass respectively through a, b, c, where $a'b'c'$ is obtained by turning abc through an angle θ about S. Thus the basal triangle of Π'' is determined as the triangle common to three systems of the type L.

We may notice one other result in this connection. The κ-pedal lines of p in regard to members of any pedo-parallel pencil contained in L must both pass through Q and be parallel, and must therefore be identical. Now it is clear from 5·21 that the only point which has the same κ-pedal line in regard to members of a pedo-parallel pencil is the focus of the parabola of 5·21, that is, the point λ such that $ij\lambda$ is a member of the pencil. Hence ijp belongs to every pedo-parallel pencil of L, so that ijp is apolar to the apolar triangle $\alpha\beta\gamma$ of L, or p is the Euler point of $\alpha\beta\gamma$. We have therefore a property of the Euler point:

7·25 *The κ-pedal line of the Euler point p of $\alpha\beta\gamma$, in regard to all triangles apolar to $\alpha\beta\gamma$, passes through a fixed point Q; Q* is the point which divides pO' in the ratio $2 - \kappa : \kappa$, where O' is the inverse of the orthocentre of $\alpha\beta\gamma$ in the second fundamental circle (5·25).

(33) *The structure of the manifold of Pedo-Similar Forms.*

A pedo-similar form being one which is complete and closed at ij in the manner of the pedal form, it can by 2·1 be written in the form

7·26 $$\Pi''(x, y) = (iy)(jy)(cx)^3 + (ix)(jx)(ax)(by)^2,$$

where from the closure property of Π'' it follows that $(ax)(by)^2$ is an auto-form. It was seen in (6) that the shape 7·26 for Π'' was not unique, but that $(cx)^3$ and $(ax)(by)^2$ can be simultaneously

Digitized by Google

replaced by $(cx)^3 - (ix)(jx)(\lambda x)$ and $(ax)(by)^2 + (iy)(jy)(\lambda x)$ respectively. Thus for a given Π'', the possible forms $(cx)^3$ are members of a pedo-parallel system L_1, and the possible forms $(ax)(by)^2$ are auto-forms of members of a second pedo-parallel system L_2. Now it is clear from the form 7·26 that L_1 is the concurrency system of Π'': hence by 7·3, if θ be the deviation of Π'', L_1 makes the pedal angle $-\frac{3}{2}\theta$ with the basal triangle of Π'. Also from the form 7·26, the auto-chord of x in regard to $(ax)(bx)^2$ is clearly parallel to the chord which corresponds to x in Π''. It follows therefore that $(ax)(bx)^2$ makes the pedal angle $-\frac{1}{2}\theta$ with the basal triangle. Hence we deduce that

7·27 *If θ be the pedal angle between the pedo-parallel systems L_1, L_2, the deviation of Π'' is either θ or $\theta + \pi$.* In particular, the necessary and sufficient condition that the form Π'' given by 7·26 should be semi-pedal, is the pedo-parallelism of $(cx)^3$ and $(ax)(bx)^2$.

It was seen in (6) that the representation 7·26 could be rendered unique by choosing $(cx)^3$ to be apolar to $(ix)(jx)$, that is, to be of the form $p_1(ix)^3 + p_2(jx)^3$ representing an equilateral triangle. If we also write the auto-form $(ax)(by)^2$ in the shape

$$(ix)(iy)(Ay) + (jx)(jy)(By) \quad (3{\cdot}7),$$

we have the unique representation of Π'' in the form:

7·28

$$\Pi''(x, y) = (iy)(jy)\{p_1(ix)^3 + p_2(jx)^3\} + (ix)(jx)\{(ix)(iy)(Ay) + (jx)(jy)(By)\}.$$

We shall call $p_1, p_2, A_1, A_2, B_1, B_2$ the six coordinates of the form Π''. To find the measure-deviation of a form in terms of its coordinates, we first find the coordinates of a form $\Pi''(x, y)$ given by its basal triangle $\alpha\beta\gamma$ and its measure-deviation (κ, θ). Writing as before $x' = \lambda(jx)i + \mu(ix)j$, so that $\lambda + e^{\sqrt{-1}\theta}\mu = 0$, and

$$(\alpha x)(\beta x)(\gamma x) \equiv (ix)^2(Px) + (jx)^2(Qx),$$

we easily find from 6·12 that

$$\begin{aligned}\Pi''(x, y) &= \Pi'(x', y)\\ &= (ij)^2\{(ix')(jx')[(ix')(iy)(Py) + (jx')(jy)(Qy)]\\ &\qquad - \kappa(iy)(jy)[(ix')^2(Px') + (jx')^2(Qx')]\}\\ &= (ij)^4\{(iy)(jy)[p_1(ix)^3 + p_2(jx)^3]\\ &\qquad + (ix)(jx)[(ix)(iy)(Ay) + (jx)(jy)(By)]\},\end{aligned}$$

Digitized by Google

where

7·29
$$\begin{cases} p_1 = -\kappa\mu^3 (Pj), \\ p_2 = -\kappa\lambda^3 (Qi), \\ A = -\lambda\mu^2 \{2(ij)P + \kappa(Pi)j\}, \\ B = \lambda^2\mu \{2(ij)Q - \kappa(Qj)i\}. \end{cases}$$

On eliminating P, Q from these, we obtain the following equations for the measure and deviation of the form Π'' with coordinates $p_1, p_2, A_1, A_2, B_1, B_2$:

7·30
$$\begin{cases} U_\theta \equiv \lambda^2 p_1 (Bi) + \mu^2 p_2 (Aj) = 0, \text{ where } \lambda + e^{\sqrt{-1}\theta}\mu = 0, \\ V_\kappa \equiv \kappa^2 (Aj)(Bi) + 4(ij)^2 p_1 p_2 = 0. \end{cases}$$

We now represent the form Π'' with the coordinates $p_1, p_2, A_1, A_2, B_1, B_2$ by a point Π'' with the same coordinates in a five-dimensional space Σ_5 which corresponds to the pedo-similar manifold. The forms with a given measure κ or a given deviation θ will then correspond to two families of quadrics U, V, whose equations are given by 7·30. We see that the general quadric U_θ, as well as the general quadric V_κ, is singular, with the same line of vertices L, given by the equations $p_1 = p_2 = (Aj) = (Bi) = 0$. The points of L correspond to what we may call *the singular pedo-similar forms*, namely those of the shape $(ix)(jx)(tx)(iy)(jy)$, where t may be arbitrary.

Since U_θ and V_κ are cones with the line-vertex L, they have two distinct systems of generating three-dimensional regions passing through L. The quadrics U have clearly the four common generating regions:

$$\gamma_1 : p_1 = p_2 = 0, \qquad \gamma_3 : p_1 = (Aj) = 0,$$
$$\gamma_2 : (Aj) = (Bi) = 0, \qquad \gamma_4 : p_2 = (Bi) = 0.$$

The quadrics V have similarly the four common generating regions:

$$\gamma_5 : p_1 = (Bi) = 0, \qquad \gamma_3 : p_1 = (Aj) = 0,$$
$$\gamma_6 : p_2 = (Aj) = 0, \qquad \gamma_4 : p_2 = (Bi) = 0.$$

The two families U, V have thus in common the two generating regions γ_3, γ_4; the system of generating regions of U or V to which γ_3, γ_4 belong will be called *the secondary system*, and the other system, *the primary system*.

The general primary generating regions of U_θ, V_κ are easily seen from 7·30 to be given by the respective equations:

$$p_1 = t\mu^2 (Aj), \qquad p_2 = -t\lambda^2 (Bi);$$
$$p_1 = t\kappa^2 (Aj), \qquad (Bi) = -4(ij)^2 t p_2.$$

Digitized by Google

From the expressions 7·30 for the measure and deviation, it follows that κ and θ are respectively constant on these generating regions. Hence we see that

7·31 *The primary generating regions of U_θ or V_κ represent families of forms with constant measure-deviation.*

In particular, the primary generating regions γ_1, γ_2 of U represent respectively the products of $(ix)(jx)$ with auto-forms, and the forms of rank one. Since U_θ and V_κ have the secondary generating regions γ_3, γ_4 in common, they must also have two primary generating regions in common; remembering the ambiguity in the measure-deviation of a pedo-similar form (7·9), it clearly follows from 7·31 that these represent the two families of forms whose measure-deviations are of the shape

$$\{(\kappa, \theta) \text{ or } (-\kappa, \theta+\pi)\}, \quad \{(\kappa, \theta+\pi) \text{ or } (-\kappa, \theta)\}.$$

If in obtaining equations 7·29, we replace the basal triangle by any other triangle pedo-parallel to it, this means that P, Q in 7·29 are increased by arbitrary scalar multiples of j, i respectively. It is clear that the following relations hold between the coordinates in 7·29, independently of such increase:

$$7\cdot32 \qquad \frac{p_1}{p_2} = \frac{\mu^3 (Pj)}{\lambda^3 (Qi)}; \quad \frac{(Aj)}{(Bi)} = \frac{-\mu}{\lambda} \frac{(Pj)}{(Qi)}.$$

If θ is given, 7·32 is satisfied by all forms of deviation θ, whose basal triangles belong to a pedo-parallel system. But, obviously, equations of the form 7·32 represent secondary generating regions of U_θ. Hence

7·33 *The general secondary generating region of U_θ represents the family of pedo-similar forms of deviation θ, whose basal triangles are members of a pedo-parallel system.*

The general secondary generating region of V_κ is given by

$$p_1 = t\kappa^2 (Bi); \quad (Aj) = -4 (ij)^2 t p_2,$$

where t is fixed for each generating region. Combining this with 7·29, we obtain

$$\mu^2 (Pj) + 2 (ij) t\kappa\lambda^2 (Qi) = 0,$$

or

$$1 + e^{2\sqrt{-1}\theta} . 2 (ij) . t\kappa . \frac{(Qi)}{(Pj)} = 0,$$

where θ is the deviation. But $\frac{1}{2\sqrt{-1}} \log \frac{(Qi)}{(Pj)}$ is easily seen to differ only by a constant from the pedal angle between an arbitrary fixed triangle and the basal triangle given by

$$(ix)^2 (Px) + (jx)^2 (Qx).$$

Digitized by Google

Thus the meaning of the above equation is that the sum of the increments of the deviation and of the pedal angle which the basal triangle makes with a fixed triangle is zero. Hence

7·34 *The general secondary generating region of V_κ represents a family of pedo-similar forms whose measure is κ, and whose basal triangles are quite arbitrary. The sum of the deviation of any form of the family and the pedal angle which its basal triangle makes with a given triangle is a constant.*

The pencil of defective polar forms of equilateral triangles corresponds to a line M; the quadric U_θ which contains M is none other than the semi-pedal quadric U_0. The semi-pedal pencils, namely those of the type 6·1, correspond to those lines on U_0 which meet γ_1, γ_2 and M.

§ 8. Extension to Space of n Dimensions.

(34) *The Pedal Correspondence on the Wallace curve.*

A norm curve in Euclidean space of n dimensions, which has the property that the feet of the perpendiculars from any point on it to the faces of any inscribed simplex lie on a prime*, may be called a Wallace curve. It has been shewn by H. W. Richmond that the Wallace curves are those norm curves which intersect the prime at infinity in the vertices of a circumscribed simplex of the Absolute†.

Let C be a Wallace curve cutting the prime at infinity in $i_1, i_2 \dots i_n$, and $\alpha_0, \alpha_1 \dots \alpha_n$ a simplex inscribed in C. Let the pedal prime of a point x on C—namely, the prime which contains the feet of the perpendiculars from x on the faces of $\alpha_0, \alpha_1 \dots \alpha_n$—cut C in y. It is easily seen that the pedal correspondence (x, y) is an $(n+1, n)$ correspondence Π. Also, the perpendicular from any of the points i on any face of the simplex lies wholly at infinity, so that the pedal prime of any i is the prime $i_1, i_2 \dots i_n$. Thus $i_1, i_2 \dots i_n$ is a complete (but not a closed) set of the pedal correspondence. Thus

8·1 *The pedal correspondence Π in n dimensions, as defined thus, is—in contrast to the two-dimensional case 6·2—complete, but not closed at infinity.*

By 2·4, the existence of the complete n-set reduces the rank of Π to three. Hence

8·2 *The pedal correspondence obtains no increase in rank through the increase in the number of dimensions.*

* The word "prime" has been suggested by Professor Baker for denoting a flat $(n-1)$-dimensional space situated in a flat space of n dimensions.

† *Proc. Camb. Phil. Soc.* xxii (1925), pp. 34—41.

Digitized by Google

From the fact that the rank is three, it follows that all pedal primes could be expressed as linear combinations of three of them. This shews that

8·3 *Every pedal prime passes through a fixed region L_{n-3} of $n-3$ dimensions situated wholly at infinity. In particular, for the three-dimensional case, every pedal plane is parallel to a fixed direction.*

The pedal envelope has the prime at infinity for an n-ple prime, the corresponding regions of contact being the $(n-2)$-dimensional regions which join L_{n-3} to $i_1, i_2, \ldots i_n$. Since the envelope is of class $n+1$ (cf. (8)), and has the prime at infinity for an n-ple prime, it follows that if any pedal prime P cuts the prime at infinity in the region L_1 (containing L), there is a homographic correspondence between P and L_1. Since there are two regions L_1, containing L, which touch the Absolute, it follows that there are two points $j_1 j_2$ on the Wallace curve whose pedal primes touch the Absolute. There is a faint analogy between $j_1 j_2$ and the circular points of the two-dimensional case. Thus the analogue of the angle-property would be:

8·4 If a homography on the curve with the fixed points $j_1 j_2$ carries p to p', then the pedal primes of p, p' make a constant angle with one another for all positions of p.

It is easy to see also that the concurrency system of the pedal form Π is a linear ∞^2-system containing the singular pencil $(i_1 x)(i_2 x)\ldots(i_n x)(\lambda x)$. Such systems are the analogues of "pedo-parallel" systems. Every such system contains a unique "equilateral" member, namely a member which is apolar to $(i_1 x)\ldots(i_n x)$. If two such systems contain the respective "equilateral" members

$$\lambda_1 (i_1 x)^{n+1} + \lambda_2 (i_2 x)^{n+1} + \ldots + \lambda_n (i_n x)^{n+1},$$

$$\mu_1 (i_1 x)^{n+1} + \mu_2 (i_2 x)^{n+1} + \ldots + \mu_n (i_n x)^{n+1},$$

the analogue of the "pedal angle" between the two systems will be a series of angles ϕ_{rs}, which are defined in terms of the ratios

$$\frac{\mu_1}{\lambda_1}, \frac{\mu_2}{\lambda_2} \cdots \frac{\mu_n}{\lambda_n},$$

by means of the equations

$$\phi_{rs} = \frac{1}{2\sqrt{-1}} \log \frac{\mu_r \lambda_s}{\mu_s \lambda_r},$$

so that $$\phi_{rs} = -\phi_{sr}, \quad \phi_{rr} = 0 \pmod{\pi}.$$

Digitized by Google

Lines and Planes in a Metrical Space. By W. HUNTER, B.A., Trinity College. (Communicated by Mr S. POLLARD.)

[*Received* 2 December 1926, *read* 31 January 1927.]

§ 1. *"Stretches" joining Points in a Metrical Space.*

If x, y, z are any points in a metrical space, then by definition

$$\delta(x, y) \leqslant \delta(z, x) + \delta(z, y).$$

Regarding x and y as fixed, we may define the set of points z for which

$$\delta(x, y) = \delta(x, z) + \delta(z, y)$$

as the "stretch joining the points xy" or more shortly the stretch xy*. Then for points z not on the stretch,

$$\delta(x, y) < \delta(x, z) + \delta(z, y).$$

The stretch xy always contains the points x and y themselves, which we may call the "end-points" of the stretch. In particular cases, it may contain no other points: e.g. with the function $\delta_0(x, y) = 1$ when $x \neq y$, each stretch contains only its end-points†.

(1) The "diameter" of a stretch is equal to the distance between its end-points. If z, z' are any two points of the stretch xy, then

$$\delta(z, z') \leqslant \delta(z, x) + \delta(x, z')$$

and

$$\delta(z, z') \leqslant \delta(z, y) + \delta(z', y),$$

and therefore by addition

$$\begin{aligned} 2\delta(z, z') &\leqslant \delta(x, z) + \delta(z, y) + \delta(x, z') + \delta(z', y) \\ &= 2\delta(x, y). \end{aligned}$$

The diameter of the stretch, the upper bound of $\delta(z, z')$, is equal to $\delta(x, y)$ [which value it attains for $z = x$, $z' = y$].

(2) If z lies on the stretch xy, then the stretch xy contains both the stretch xz and the stretch zy.

For let t be any point of the stretch xz. Then

$$\begin{aligned} \delta(x, y) &= \delta(x, z) + \delta(z, y) \\ &= \delta(x, t) + \delta(t, z) + \delta(z, y) \\ &\geqslant \delta(x, t) + \delta(t, y) \\ &\geqslant \delta(x, y). \end{aligned}$$

* See Menger, "Ueber geodätische Linien in allgemeinen metrischen Räumen," *Amsterdam Proceedings*, Vol. XXIX (1926), p. 166.

† If each stretch contains at least one point in addition to its end-points, the space is said to be "convex." See Menger, *op. cit.*

Digitized by Google

Hence the equality sign must hold throughout; t lies on the stretch xy, and incidentally we have proved that z lies on the stretch ty. Similarly we may prove that the stretch xy contains the stretch zy and that, if w is any point of the stretch zy, then z lies on the stretch xw. The two stretches xz, zy can obviously have no point in common except the end-point z.

Examples will shew (see § 2) that the stretch xy is not necessarily the sum of the two stretches xz, zy defined by any point of it. If, therefore, it has this property we call it a "linear" stretch.

(3) The necessary and sufficient condition for a stretch to be linear is that z, z' being any two points of the stretch, then

$$\delta(z, z') = |\delta(x, z) - \delta(x, z')| = |\delta(z, y) - \delta(z', y)|,$$

i.e. that the stretch should be applicable on a set of points lying on a euclidean straight line.

The condition is necessary; for if the stretch is linear, then either z' lies on the stretch xz and

$$\delta(z, z') = \delta(x, z) - \delta(x, z'),$$

or else z' lies on the stretch zy and

$$\delta(z, z') = \delta(z, y) - \delta(z', y).$$

It is evidently sufficient; for z being any point of the stretch divides it into the stretches xz, xy, and we can deduce from the given condition that any other point z' must belong to one or other of them: if

$$\delta(z, z') = \delta(x, z) - \delta(x, z'),$$

then
$$\delta(x, z) = \delta(x, z') + \delta(z', z),$$

and z' lies on the stretch xz; if

$$\delta(z, z') = \delta(x, z') - \delta(x, z),$$

then
$$\delta(z, z') = \delta(y, z) - \delta(y, z'),$$

and therefore
$$\delta(y, z) = \delta(y, z') + \delta(z', z),$$

and z' lies on the stretch yz. The condition being satisfied, $\delta(x, z)$ regarded as a function of z can take each of its possible values once only, for if z, z' are two points for which

$$\delta(x, z) = \delta(x, z'),$$

then
$$\delta(z, z') = |\delta(x, z) - \delta(x, z')| = 0,$$

and z and z' coincide. We can then establish a (1, 1) relation between the points z of the stretch xy and points of the interval $[0, \delta(x, y)]$ of, say, the x-axis in the euclidean space, by letting

Digitized by Google

z correspond to the point of abscissa $\delta(x, z)$ and this is an application for

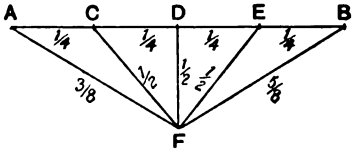
$$\delta(z_1, z_2) = |\delta(x, z_1) - \delta(x, z_2)|.$$

= distance between corresponding points of the interval on the euclidean line.

§ 2. *Examples of Non-linear Stretches.*

(i) Let Σ consist of the six points A, B, C, D, E and F, and let

$$\delta(A, C) = \delta(C, D) = \delta(D, E) = \delta(E, B) = \tfrac{1}{4},$$
$$\delta(A, F) = \tfrac{3}{8},$$
$$\delta(A, D) = \delta(C, E) = \delta(D, B) = \delta(C, F) = \delta(D, F) = \delta(E, F) = \tfrac{1}{2},$$
$$\delta(F, B) = \tfrac{5}{8},$$
$$\delta(A, E) = \delta(C, B) = \tfrac{3}{4},$$
$$\delta(A, B) = 1.$$



This is easily seen to be a metrical space. The stretch AB consists of all six points and $\delta(A, X)$ for points X of the stretch takes each of its possible values once only. The stretch is not linear, for the point F lies neither on the stretch AD nor on the stretch DB.

(ii) Let Σ consist of the points $x \equiv (x_1, x_2, \ldots, x_n)$, and let

$$\delta(x, y) = |x_1 - y_1| + |x_2 - y_2| + \ldots + |x_n - y_n|.$$

It may easily be verified that this is a metrical space. The stretch xy consists of all the points of the "rectangular solid" with faces parallel to the co-ordinate planes, having x and y as opposite corners; in 2 dimensions, for example, the stretch consists of all the points of a rectangle and is easily seen to be non-linear.

(iii) Let the points of Σ be the bounded functions $f(x)$ defined, say, for the interval $0 \leqslant x \leqslant 1$, and let the distance function be

$$\delta(f_1, f_2) = \|f_1(x) - f_2(x)\|$$
= upper bound of $|f_1(x) - f_2(x)|$ for $0 \leqslant x \leqslant 1$.

Digitized by Google

Since the distance depends only on the difference of the two functions, it is sufficient to consider the stretch joining the point 0 to the point $f(x)$. Let $\|f(x)\| = D$ and let d be any number such that $0 \leqslant d \leqslant D$.

If $$f(x) + d \geqslant \lambda(x) \geqslant f(x) - d \quad \ldots\ldots\ldots\ldots\ldots\ldots (1),$$

and $$D - d \geqslant \lambda(x) \geqslant -(D - d) \quad \ldots\ldots\ldots\ldots\ldots\ldots (2),$$

then $$\|\lambda(x)\| \leqslant D - d \text{ and } \|f(x) - \lambda(x)\| \leqslant d \ldots\ldots\ldots (3).$$

Since $$D \geqslant f(x) \geqslant -D,$$

$$f(x) - d \leqslant D - d \text{ and } f(x) + d \geqslant -(D - d),$$

and hence a function $\lambda(x)$ can always be found to satisfy (1) and (2) simultaneously. We shew that all functions $\lambda(x)$ which satisfy these inequalities lie on the stretch joining 0 to $f(x)$. D = upper bound of $|f(x)|$, and therefore given ϵ, there exists a point x_0 for which either $f(x_0) > D - \epsilon$, or else $f(x_0) < -D + \epsilon$. If the former, then

$$\lambda(x_0) \geqslant f(x_0) - d > (D - d) - \epsilon,$$

and therefore $$|\lambda(x_0)| > (D - d) - \epsilon.$$

Therefore $$\|\lambda(x)\| \geqslant D - d,$$

and therefore, by (3),

$$\|\lambda(x)\| = D - d.$$

Similarly, if $f(x_0) < -D + \epsilon$, we find again that

$$\|\lambda(x)\| = D - d.$$

Also $$\|\lambda(x)\| + \|f(x) - \lambda(x)\| \geqslant \|f(x)\| = D,$$

and therefore putting in the value for $\|\lambda(x)\|$ we find

$$\|f(x) - \lambda(x)\| \geqslant d,$$

and therefore, by (3),

$$\|f(x) - \lambda(x)\| = d.$$

The stretch can only be linear if $\delta(0, \lambda(x))$ takes each of its possible values once only. Hence since, in general, more than one function $\lambda(x)$ can be found to satisfy inequalities (1) and (2) for a given value of d, the stretches in function space are not in general linear.

If for all values of x, $|f(x)| = D$, then $f(x) - d = D - d$ or $f(x) + d = -(D - d)$ for all values of x and in this case inequalities (1) and (2) define a unique function $\lambda(x)$ always equal either to $D - d$ or to $-(D - d)$. Hence there exist linear stretches, viz. those joining points $f_1(x), f_2(x)$ for which $|f_1(x) - f_2(x)|$ is constant.

Suppose, for instance, that

$f_2(x) = f_1(x) + D$ for points x of a set E contained in $[0, 1]$

$= f_1(x) - D$ for points x of the complementary of E.

Digitized by Google

It can easily be shewn, from the above argument, that the stretch joining $f_1(x)$ and $f_2(x)$ consists of functions $\lambda(x)$ such that

$$\lambda(x) = f_1(x) + d$$

for points of E,

$$= f_1(x) - d$$

for points of the complementary of E, for all values of d satisfying $0 \leqslant d \leqslant D$. This stretch is evidently linear.

§ 3. *The Extended Stretch defined by two Points.*

Consider, in any metrical space, the function $D(x, y, z)$ defined as follows. Let

$$2s(x, y, z) = \delta(y, z) + \delta(z, x) + \delta(x, y),$$

and put

$$D(x, y, z) = +\sqrt{s_{xyz}(s_{xyz} - \delta_{yz})(s_{xyz} - \delta_{zx})(s_{xyz} - \delta_{xy})}.$$

By the axioms of a metrical space, each of the four quantities s, $s - \delta$ is $\geqslant 0$ and hence $D(x, y, z)$ exists and is $\geqslant 0$ for all points x, y, z.

Now $D(x, y, z) = 0$ only in the following cases:

(i) $s(x, y, z) = 0$, in which case x, y, and z must all coincide.

(ii) $s(x, y, z) = \delta(y, z)$, i.e. $\delta(y, z) = \delta(z, x) + \delta(x, y)$, so that x lies on the stretch yz.

(iii) $s(x, y, z) = \delta(z, x)$, i.e. y lies on the stretch zx.

(iv) $s(x, y, z) = \delta(x, y)$, i.e. z lies on the stretch xy.

These last three possibilities are evidently mutually exclusive, i.e. if

$$\delta(y, z) = \delta(z, x) + \delta(x, y),$$

then $$s(x, y, z) = \delta(y, z) = \delta(z, x) + \delta(x, y),$$

and therefore $$s(x, y, z) > \delta(x, y),$$

and $$s(x, y, z) > \delta(z, x).$$

If, then, x and y are any two given (and distinct) points, the equation $D(x, y, z) = 0$ defines a set of points z which we may call the "extended stretch xy."

If z belongs to the stretch joining x and y, it also belongs to the extended stretch. If z belongs to the extended stretch, but not to the stretch joining x and y, then of the two stretches joining z and x, z and y, only one necessarily belongs to the extended stretch; thus if z is such that y lies on the stretch xz, then the stretch yz is contained in the extended stretch, but not necessarily the stretch xz. [If y lies on xz and t is any point of yz, then by § 1, (2), t belongs to xz and y belongs to xt and therefore

Digitized by Google

t belongs to the extended stretch xy. But in Ex. (i), § 2, the extended stretch AD consists of the five points A, B, C, D, E and therefore does not contain the stretch AB.]

The particular function $D(x, y, z)$ here used to define the extended stretch is evidently not the only function which can be used for this purpose. If $f(\lambda)$ is any function of the real variable λ such that $f(\lambda) = 0$ if and only if $\lambda = 0$ and $f(\lambda) > 0$ for $\lambda > 0$, then we can take

$$D(x, y, z) = f(s)\, f(s - \delta_{yz})\, f(s - \delta_{zx})\, f(s - \delta_{xy}).$$

Extending the result (3) of § 1, we may say that the extended stretch is linear if it is applicable on a euclidean straight line.

Triangulated Spaces.

§ 4. *Definition.*

The structure of a metrical space is defined by the aid of a function $\delta(x, y)$ of its points taken in pairs, and this leads to the special characteristic we have just considered, that any pair of points defines a stretch of points which in special cases has the properties associated with lines. If we endeavour to generalize this result, we might be tempted to define what we may call a "triangulated" space, as follows: In the set Σ we define a function $\Delta(x, y, z)$ of all possible triads of points, such that

(i) $\Delta(x, y, z) \geqslant 0$ and $= 0$, when any two of the points coincide.

(ii) $\Delta(x, y, z)$ is independent of any order in the points x, y, z.

(iii) $\Delta(x, y, z) \leqslant \Delta(w, y, z) + \Delta(w, z, x) + \Delta(w, x, y)$.

These conditions are self-consistent; they may be satisfied by taking the function Δ to be 1 when x, y, z are all distinct, and 0 when any two of them coincide. They are not, however, sufficient to define a manageable function Δ.

Regarding x, y, z as given, we define the "triangle xyz" as the set of points w for which

$$\Delta(x, y, z) = \Delta(w, y, z) + \Delta(w, z, x) + \Delta(w, x, y).$$

If w is a point of the triangle xyz, then this triangle does not necessarily contain the points of the triangles wyz, wzx, wxy.

Example. Let Σ consist of the five points A, B, C, D, E and define

$$\Delta(A, B, C) = 1,$$
$$\Delta(A, B, D) = \Delta(B, C, D) = \Delta(C, A, D) = \tfrac{1}{3},$$
$$\Delta(A, B, E) = \Delta(B, D, E) = \Delta(D, A, E) = \tfrac{1}{3},$$
$$\Delta(E, B, C) = \Delta(E, C, A) = \tfrac{1}{2},$$
$$\Delta(C, D, E) = \tfrac{2}{3}.$$

Digitized by Google

Then D is a point of the triangle ABC, and E a point of the triangle ABD, but E does not belong to the triangle ABC.

§ 5. *A Metrical Space as a Triangulated Space.*

The function $D(x, y, z)$ defined in § 3 evidently satisfies conditions (i) and (ii) of § 4. If it should also satisfy condition (iii) we should be able to use it to define the metrical space as a triangulated space. We shall see, however, that $D(x, y, z)$ does not necessarily satisfy this condition.

For brevity write

$$\delta(y, z) = a, \quad \delta(z, x) = b, \quad \delta(x, y) = c,$$

and

$$\delta(w, x) = l, \quad \delta(w, y) = m, \quad \delta(w, z) = n.$$

Since the space is metrical, each of the quantities

$$m + n + a, \quad m + n - a, \quad m - n + a, \quad -m + n + a$$

is $\geqslant 0$

and hence the equations

$$2\sqrt{mn}\sin\frac{\alpha}{2} = \sqrt{(m - n + a)(-m + n + a)},$$

$$2\sqrt{mn}\cos\frac{\alpha}{2} = \sqrt{(m + n + a)(m + n - a)}$$

define uniquely a real angle α when $0 \leqslant \alpha \leqslant \pi$. We note that $\alpha = \pi$ if w lies on the stretch joining y and z, and $\alpha = 0$ if w lies elsewhere on the extended stretch yz.

We easily deduce

$$a^2 = m^2 + n^2 - 2mn\cos\alpha,$$

$$D(w, y, z) = \tfrac{1}{2}mn\sin\alpha.$$

Defining the angles β, γ in a similar manner, we have

$$D(w, z, x) = \tfrac{1}{2}nl\sin\beta$$

and

$$D(w, x, y) = \tfrac{1}{2}lm\sin\gamma,$$

$$\begin{aligned}
&D(x, y, z)\\
&= \tfrac{1}{4}\sqrt{(a + b + c)(a + b - c)(a - b + c)(-a + b + c)}\\
&= \tfrac{1}{4}\sqrt{2\Sigma b^2c^2 - \Sigma a^4}\\
&= \tfrac{1}{4}\sqrt{\Sigma[2(n^2 + l^2 - 2nl\cos\beta)(l^2 + m^2 - 2lm\cos\gamma) - (m^2 + n^2 - 2mn\cos\alpha)^2]}\\
&= \tfrac{1}{2}\sqrt{\Sigma m^2n^2\sin^2\alpha + 2lmn\Sigma l(\cos\beta\cos\gamma - \cos\alpha)}\\
&= \tfrac{1}{2}\sqrt{[\Sigma mn\sin\alpha]^2 - 4lmn\sin\sigma\,\Sigma l\sin(\sigma - \alpha)},
\end{aligned}$$

Digitized by Google

where $$\sigma = \frac{\alpha+\beta+\gamma}{2}.$$

The remaining conditions to ensure that the space is metrical are evidently equivalent to the condition that the expression under the $\sqrt{\ }$ sign should be positive; i.e. the space is metrical if and only if

$$0 \leqslant \alpha \leqslant \pi, \quad 0 \leqslant \beta \leqslant \pi, \quad 0 \leqslant \gamma \leqslant \pi,$$

$$[\Sigma mn \sin \alpha]^2 - 4lmn \sin \sigma \, \Sigma l \sin (\sigma - \alpha) \geqslant 0 \ldots\ldots\ldots\ldots (\mathrm{M})$$

Since $$D(w, y, z) + D(w, z, x) + D(w, x, y) = \tfrac{1}{2} \Sigma mn \sin \alpha,$$

$$D(x, y, z) \leqslant D(w, y, z) + D(w, z, x) + D(w, x, y) \ldots (\mathrm{iii})$$

if and only if $$4lmn \sin \sigma \, \Sigma l \sin (\sigma - \alpha) \geqslant 0.$$

(a) Now if $\sigma > \pi$, then $\sigma - \alpha > 0$ ($\alpha \leqslant \pi$), and also

$$\sigma - \beta > 0, \ \sigma - \gamma > 0; \ \sigma = \frac{\alpha+\beta+\gamma}{2} \leqslant \frac{3\pi}{2},$$

and therefore since

$$\sigma = \sigma - \alpha + \sigma - \beta + \sigma - \gamma,$$

one at least of the angles $\sigma - \alpha$, $\sigma - \beta$, $\sigma - \gamma$ must be different from π,

$$\sigma - \alpha = \frac{\beta+\gamma-\alpha}{2} \leqslant \pi \text{ etc.}$$

Hence $$\sin \sigma < 0$$

and $$\sin (\sigma - \alpha), \ \sin (\sigma - \beta), \ \sin (\sigma - \gamma)$$

cannot be negative or all zero. Therefore

$$4lmn \sin \sigma \, \Sigma l \sin (\sigma - \alpha) < 0.$$

Condition (M) is certainly satisfied, but (iii) cannot be true.

If $\sigma = \pi$, then evidently

$$D(x, y, z) = D(w, y, z) + D(w, z, x) + D(w, x, y).$$

When $\sigma < \pi$, there are two possibilities to consider:

(b) $$\sigma - \alpha \geqslant 0, \quad \sigma - \beta \geqslant 0, \quad \sigma - \gamma \geqslant 0.$$

In this case, since $\sigma - \alpha + \sigma - \beta + \sigma - \gamma = \sigma < \pi$, none of the three angles can be π. They can only vanish simultaneously when $\sigma = \alpha = \beta = \gamma = 0$ and as this implies that the points x, y, z coincide, we exclude this possibility. Hence we have

$$4lmn \sin \sigma \, . \, \Sigma l \sin (\sigma - \alpha) > 0,$$

i.e. $$D(x, y, z) < D(w, z, x) + D(w, x, y) + D(w, y, z).$$

Digitized by Google

The condition (M) is certainly satisfied in this case. We have to shew that

$$\Delta^2 \equiv [\Sigma mn \sin \alpha]^2 - 4lmn \sin \sigma \, \Sigma l \sin (\sigma - \alpha) \geqslant 0.$$

Putting $mn = \xi, \quad nl = \eta, \quad lm = \zeta,$

we have $\Delta^2 = \Sigma \xi^2 \sin^2 \alpha + 2\Sigma \eta\zeta (\cos \beta \cos \gamma - \cos \alpha).$

By a well-known theorem Δ^2 can be put in the form

$$\lambda_1 X^2 + \lambda_2 Y^2 + \lambda_3 Z^2,$$

where X, Y, Z are real linear functions of ξ, η, ζ and $\lambda_1 \lambda_2 \lambda_3$ are the (real) roots of the cubic

$$\begin{vmatrix} \sin^2 \alpha - \lambda, & \cos \alpha \cos \beta - \cos \gamma, & \cos \gamma \cos \alpha - \cos \beta \\ \cos \alpha \cos \beta - \cos \gamma, & \sin^2 \beta - \lambda, & \cos \beta \cos \gamma - \cos \alpha \\ \cos \gamma \cos \alpha - \cos \beta, & \cos \beta \cos \gamma - \cos \alpha, & \sin^2 \gamma - \lambda \end{vmatrix} = 0,$$

i.e. $\lambda^3 - \lambda^2 (\sin^2 \alpha + \sin^2 \beta + \sin^2 \gamma) + 3\lambda\Omega - \Omega^2 = 0,$

where $\Omega = 1 - \cos^2 \alpha - \cos^2 \beta - \cos^2 \gamma + 2 \cos \alpha \cos \beta \cos \gamma$

$= 4 \sin \sigma \sin (\sigma - \alpha) \sin (\sigma - \beta) \sin (\sigma - \gamma)$

$\geqslant 0.$

We readily deduce that the roots of this cubic are all positive and therefore Δ^2 is essentially positive, and condition (M) is satisfied.

We may write the condition $\Omega \geqslant 0$ in the form

$$[V(x, y, z, w)]^2 \geqslant 0,$$

so that

$$144 [V(x, y, z, w)]^2$$

$$= 4l^2 m^2 n^2 \Omega$$

$$= 4l^2 m^2 n^2 - \Sigma l^2 (m^2 + n^2 - a^2)^2 + (m^2 + n^2 - a^2)(n^2 + l^2 - b^2)(l^2 + m^2 - c^2)$$

$$= -a^2 b^2 c^2 + \Sigma a^2 l^2 (b^2 + c^2 - a^2) - \Sigma a^2 (l^2 - m^2)(l^2 - n^2),$$

and we have just seen that this condition is sufficient to make

$$D(x, y, z) \leqslant D(w, y, z) + D(w, z, x) + D(w, x, y)$$

[for $\Omega \geqslant 0$ if and only if $\sigma \leqslant \pi$ and $\sigma - \alpha \geqslant 0, \sigma - \beta \geqslant 0, \sigma - \gamma \geqslant 0$; for $\Omega \geqslant 0$ if these conditions are satisfied; if one of them is not satisfied it can easily be shewn that the other three certainly are and therefore one and only one of the expressions

$$\sin \sigma, \ \sin (\sigma - \alpha), \ \sin (\sigma - \beta), \ \sin (\sigma - \gamma)$$

is $< 0,$

and therefore $\Omega < 0.$]

Digitized by Google

It can easily be seen that the expression for $144\,[V(x, y, z, w)]^2$ is unaltered by interchanging b and m, c and n, i.e. by interchanging x and w, and so also is unaltered by interchanging any two of the points x, y, z, w. Hence the above condition is sufficient to make any one of the expressions $D(x, y, z)$, $D(w, y, z)$, $D(w, z, x)$, $D(w, x, y) \leqslant$ the sum of the other three and condition (iii) of § 4 is completely satisfied.

In the special case $\sigma - \alpha = 0$, then $\sigma - \beta = \gamma$, $\sigma - \gamma = \beta$,

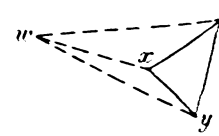
$$D(x, y, z) = \tfrac{1}{2}\sqrt{[\Sigma mn \sin\alpha]^2 - 4lmn \sin\alpha (m \sin\gamma + n \sin\beta)}$$

$$= \tfrac{1}{2}\sqrt{[mn \sin\alpha - nl \sin\beta - lm \sin\gamma]^2},$$

i.e. $$D(x, y, z) = \pm [D(w, y, z) - D(w, z, x) - D(w, x, y)],$$

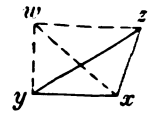
the sign being so chosen that the R.H.S. is positive, and therefore we have either

$$D(w, y, z) = D(x, y, z) + D(w, z, x) + D(w, x, y)$$



or

$$D(w, y, z) + D(x, y, z) = D(w, z, x) + D(w, x, y).$$



(c) The other alternative is: $\sigma < \pi$ and one of the angles $\sigma - \alpha$, $\sigma - \beta$, $\sigma - \gamma$ negative; say, $\alpha > \sigma$, and then $\sigma - \beta > 0$, $\sigma - \gamma > 0$. In the expression $4lmn \sin\sigma \Sigma l \sin(\sigma - \alpha)$ two of the terms are positive and one negative and hence condition (iii) is satisfied for certain values of l, m, n and not for others. Similarly we shall see that condition (M) is only satisfied for a restricted set of values of l, m, n.

We may sum up the results obtained so far as follows:

(a) $\sigma > \pi$; $\sigma - \alpha > 0$, $\sigma - \beta > 0$, $\sigma - \gamma > 0$. $D(x, y, z)$ does not satisfy condition (iii). $\Omega < 0$ and therefore $V(x, y, z, w)$ is imaginary.

(b) $\sigma \leqslant \pi$; $\sigma - \alpha \geqslant 0$, $\sigma - \beta \geqslant 0$, $\sigma - \gamma \geqslant 0$. $D(x, y, z)$ satisfies condition (iii). $\Omega \geqslant 0$ and therefore $V(x, y, z, w)$ is real.

(c) $\sigma < \pi$ and one (and then only one) of $\sigma - \alpha$, $\sigma - \beta$, $\sigma - \gamma < 0$. $D(x, y, z)$ satisfies condition (iii) only for a restricted set of points w. $\Omega < 0$ and therefore $V(x, y, z, w)$ is imaginary.

§ 6. *Metrical Spaces in which the Triangulation is Plane.*

If the metrical space Σ is such that the function $\delta(x, y)$, in addition to satisfying the three conditions for a metrical space, also satisfies, for any four points x, y, z, w, the condition (4)

$$[V(x, y, z, w)]^2 \geqslant 0,$$

Digitized by Google

where

$$144\,[V(x, y, z, w)]^2$$
$$\equiv -a^2b^2c^2 + \Sigma a^2l^2(b^2+c^2-a^2) - \Sigma a^2(l^2-m^2)(l^2-n^2),$$

we say that the triangulation is "plane."

This is equivalent to the conditions $\sigma \leqslant \pi$, $\sigma - \alpha \geqslant 0$, $\sigma - \beta \geqslant 0$ and $\sigma - \gamma \geqslant 0$ and hence $D(x, y, z)$ satisfies the conditions of § 4.

The points of the triangle xyz are defined by the condition $\sigma = \pi$. [This, as we have seen, is the only case which leads to

$$D(x, y, z) = D(w, y, z) + D(w, z, x) + D(w, x, y).]$$

We can define an extension of this triangle, as the set of points for which $V(x, y, z, w) = 0$. In this extension, we may call a point w for which $\sigma = \pi$ an "interior" point and a point for which $V(x, y, z, w) = 0$ in virtue of one of the relations $\sigma = \alpha$, $\sigma = \beta$, $\sigma = \gamma$ an "exterior" point.

If w lies on the stretch yz, then $\alpha = \pi$. Since $\sigma \leqslant \pi$, $\sigma - \alpha \geqslant 0$, therefore $\sigma = \alpha = \pi$; the point belongs to the extended triangle and may be regarded indifferently as an interior or an exterior point. If w lies elsewhere on the extended stretch yz, then $\alpha = 0$, and hence $\sigma - \beta \geqslant 0$, $\sigma - \gamma \geqslant 0$ only if $\beta = \gamma$, and therefore $\sigma = \beta = \gamma$. Then w belongs to the extended triangle and is an exterior point. Similar results hold for the stretches determined by the points zx, xy.

The stretch joining any two points of such a space is linear. Let x and y be the two points and z, z' any two points of the stretch joining them. We consider the angles α, β, γ subtended at x by the pairs of points $z'y$, yz, zz'. Since

$$\delta(z', y) = \delta(x, y) - \delta(x, z'),$$

the angle $\alpha = 0$; similarly $\beta = 0$, and therefore from $\sigma - \gamma \geqslant 0$ we deduce $\gamma = 0$, i.e.

$$\{\delta(z, z')\}^2 = \{\delta(x, z) - \delta(x, z')\}^2$$
$$\delta(z, z') = |\,\delta(x, z) - \delta(x, z')\,|,$$

and hence, by § 1 (3), the stretch joining x and y is linear.

Let X, Y, Z be any three points of Σ and W any point of the triangle determined by them, so that

$$Y\hat{W}Z = \alpha, \quad Z\hat{W}X = \beta, \quad X\hat{W}Y = \gamma,$$

the angles α, β, γ being determined by trigonometrical formulae involving the distances. Of the three distances YZ, ZX, XY, any one is not greater than the sum of the other two and we can determine a triangle in a euclidean plane, the lengths of whose sides are equal to these distances. It can easily be shewn that there exists one and only one point in the euclidean plane whose

Digitized by Google

distances from the vertices of this triangle are equal respectively to the distances WX, WY, WZ (the existence of the point is a consequence of the relation $\alpha+\beta+\gamma=2\pi$). Hence the set of four points x, y, z, w can be applied to a similar set of points lying in a euclidean plane and the angles in the two cases, being given by identical trigonometrical formulae, are equal. If, then, W is an interior point of the triangle XYZ, so that

$$Y\hat{W}Z + Z\hat{W}X + X\hat{W}Y = 2\pi,$$

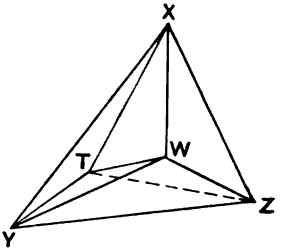
then for the angles subtended at X by YZ, ZW, WY we have the relation $Y\hat{X}Z = Y\hat{X}W + W\hat{X}Z$, and hence X is an exterior point of the extended triangle defined by WYZ. [We could have deduced this latter result directly from the equation

$$V(x, y, z, w) = 0,$$

but this does not easily lend itself to a deduction of the particular relation between the angles subtended at X.] Similar results hold for the points Y and Z.

Let now T be an interior point of, say, the triangle WXY. Then by the result just proved,

$$\begin{aligned} X\hat{Y}Z &= X\hat{Y}W + W\hat{Y}Z \\ &= X\hat{Y}T + W\hat{Y}T + W\hat{Y}Z \\ &\geqslant X\hat{Y}T + T\hat{Y}Z \\ &\geqslant X\hat{Y}Z. \end{aligned}$$



Hence the equality must hold throughout

$$X\hat{Y}Z = X\hat{Y}T + T\hat{Y}Z.$$

Similarly we prove

$$Z\hat{X}Y = Z\hat{X}T + Y\hat{X}T,$$

and from these we can deduce (as in the euclidean case, the angles being determined by identical trigonometrical formulae) that T is an interior point of the triangle XYZ.

As in the case of stretches joining two points, the triangle XYZ is not necessarily the sum of the triangles WYZ, WZX, WXY, when W is a point of the triangle XYZ (using the term triangle in the restricted sense, confining it to interior points). Let

$$YZ = ZX = XY = \sqrt{3}$$

and

$$WX = WY = WZ = 1, \quad W'X = W'Y = W'Z = 1, \quad WW' = 1.$$

Digitized by Google

These distances satisfy all the requisite conditions. We find that W and W' are both points of the triangle XYZ, but W' is exterior to each of the triangles WYZ, WZX, WXY. [E.g.

$$Y\hat{W}'Z = \frac{2\pi}{3}, \quad W\hat{W}'Y = \frac{\pi}{3}, \quad W\hat{W}'Z = \frac{\pi}{3},$$

and therefore

$$Y\hat{W}'Z + W\hat{W}'Y + W\hat{W}'Z = \frac{4\pi}{3} \neq 2\pi,$$

and similarly in the other cases.]

If the triangle XYZ is the sum of the triangles WYZ, WZX, WXY for any interior point W, then the argument used above shews that the points of the triangle XYZ are applicable on a set of points lying in a euclidean plane. This condition is evidently sufficient as well as necessary. It is for this reason that I have called the space one in which the triangulation is plane.

§ 7. *Metrical Spaces in which the Triangulation is not Plane.*

In this case, we are to consider the possibility that $\sigma < \pi$ and that one of the angles $\sigma - \alpha$, $\sigma - \beta$, $\sigma - \gamma$ is negative. Let us suppose $\alpha > \sigma$. Then $\sigma - \beta > 0$, $\sigma - \gamma > 0$.

Since $\pi \geqslant \alpha = \beta + \gamma + 2(\alpha - \sigma)$ we also have

$$\alpha - \sigma \leqslant \tfrac{1}{2}\pi, \quad \beta + \gamma < \pi.$$

We make use of the following identities

$$\sin\beta \sin\gamma = \sin(\sigma - \beta)\sin(\sigma - \gamma) + \sin\sigma \sin(\sigma - \alpha),$$

and two similar results,

$$\sin^2(\sigma - \beta) + \sin^2(\sigma - \gamma) - \sin^2\alpha = -2\cos\alpha \sin(\sigma - \beta)\sin(\sigma - \gamma) \text{ etc.},$$

$$\sin^2(\sigma - \beta) - \sin^2(\sigma - \gamma) - \sin^2\alpha = -2\cos(\sigma - \beta)\sin\sigma \sin(\sigma - \gamma) \text{ etc.},$$

$$\sin^2(\sigma - \alpha) + \sin^2\sigma - \sin^2\alpha = 2\cos\alpha \sin\sigma \sin(\sigma - \alpha) \text{ etc.}$$

We consider first the condition (M) which, putting $mn = \xi$, $nl = \eta$, $lm = \zeta$, we write

$$\Delta^2 = \Sigma\xi^2 \sin^2\alpha + 2\Sigma\eta\zeta(\cos\beta\cos\gamma - \cos\alpha) \geqslant 0.$$

Digitized by Google

Now

$$\Delta^2 = \Big[\xi \sin\alpha - \frac{1}{\sin\alpha}\{\eta(\sin\sigma\sin\overline{\sigma-\gamma} + \sin\overline{\alpha-\sigma}\sin\overline{\sigma-\beta}) + \zeta(\sin\sigma\sin\sigma-\beta + \sin\alpha-\sigma\sin\overline{\sigma-\gamma})\}\Big]^2$$
$$+ \eta^2\sin^2\beta + \zeta^2\sin^2\gamma$$
$$+ 2\eta\zeta\{\sin\overline{\sigma-\beta}\sin\overline{\sigma-\gamma} + \sin\sigma\sin\overline{\alpha-\sigma}\}$$
$$- \frac{1}{\sin^2\alpha}\{\eta(\sin\sigma\sin\sigma-\gamma + \sin\alpha-\sigma\sin\overline{\sigma-\beta}) + \zeta(\sin\sigma\sin\overline{\sigma-\beta} + \sin\alpha-\sigma\sin\overline{\sigma-\gamma})\}^2.$$

In this we have assumed that $\sin\alpha \neq 0$. We consider later the case $\alpha = \pi$. The terms not containing ξ can be shewn to reduce to

$$-\frac{\Omega}{\sin^2\alpha}\{\eta^2 - 2\eta\zeta\cos\alpha + \zeta^2\} = -\frac{\Omega}{\sin^2\alpha}l^2(m^2+n^2-2mn\cos\alpha)$$
$$= -\frac{\Omega}{\sin^2\alpha}a^2l^2,$$

where

$$\Omega = -4\sin\sigma\sin(\sigma-\alpha)\sin(\sigma-\beta)\sin(\sigma-\gamma) \geqslant 0.$$

Hence the condition $\Delta^2 \geqslant 0$ gives rise to two possibilities

$$\xi\sin^2\alpha \geqslant \eta\{\sin\sigma\sin(\sigma-\gamma) + \sin(\alpha-\sigma)\sin(\sigma-\beta)\} + \zeta\{\sin\sigma\sin(\sigma-\beta) + \sin(\alpha-\sigma)\sin(\sigma-\gamma)\} + \sqrt{\Omega}.al,$$
$$\xi\sin^2\alpha \leqslant \eta\{\ldots\} + \zeta\{\ldots\} - \sqrt{\Omega}.al.$$

If the former condition holds, then we have

$$\sin\alpha[\xi\sin\alpha - \eta\sin\beta - \zeta\sin\gamma]$$
$$\geqslant 2\eta\sin(\alpha-\sigma)\sin(\sigma-\beta) + 2\zeta\sin(\alpha-\sigma)\sin(\sigma-\gamma) + \sqrt{\Omega}.al$$
$$\geqslant 0,$$

all the terms being positive. Further

$$\Delta^2 \equiv [\xi\sin\alpha - \eta\sin\beta - \zeta\sin\gamma]^2 - 4\sin(\alpha-\sigma)\{-\eta\zeta\sin\sigma + \zeta\xi\sin(\sigma-\gamma) + \xi\eta\sin(\sigma-\beta)\},$$

and

$$\sin^2\alpha\{-\eta\zeta\sin\sigma + \zeta\xi\sin(\sigma-\gamma) + \xi\eta\sin(\sigma-\beta)\}$$
$$\geqslant -\eta\zeta\sin\sigma\sin^2\alpha + \sin(\alpha-\sigma)\{\eta\sin(\sigma-\beta) + \zeta\sin(\sigma-\gamma)\}^2 + \sin\sigma\sin(\sigma-\beta)\sin(\sigma-\gamma)\{\eta^2+\zeta^2\} + \eta\zeta\sin\sigma\{\sin^2(\sigma-\beta) + \sin^2(\sigma-\gamma)\} + \sqrt{\Omega}\sqrt{\eta^2 - 2\eta\zeta\cos\alpha + \zeta^2}\{\eta\sin(\sigma-\beta) + \zeta\sin(\sigma-\gamma)\}$$
$$= [\sqrt{\sin(\alpha-\sigma)}\{\eta\sin(\sigma-\beta) + \zeta\sin(\sigma-\gamma)\} + \sqrt{\sin\sigma\sin(\sigma-\beta)\sin(\sigma-\gamma)}\sqrt{\eta^2 - 2\eta\zeta\cos\alpha + \zeta^2}]^2$$
$$\geqslant 0.$$

Digitized by Google

Therefore

$$D(x, y, z) = \tfrac{1}{2}\Delta \leqslant \tfrac{1}{2}[\xi \sin\alpha - \eta \sin\beta - \zeta \sin\gamma],$$

i.e. $$D(w, y, z) \geqslant D(x, y, z) + D(w, z, x) + D(w, x, y).$$

In this case, therefore, the function D does not satisfy the required condition.

Hence we need only consider the case where

$$\xi \sin^2\alpha \leqslant Q \equiv \sin\sigma\{\eta \sin(\sigma-\gamma) + \zeta \sin(\sigma-\beta)\} + \sin(\alpha-\sigma)\{\eta \sin(\sigma-\beta) + \zeta \sin(\sigma-\gamma)\} - \sqrt{\Omega}.al.$$

The condition

$$D(x, y, z) \leqslant D(w, y, z) + D(w, z, x) + D(w, x, y),$$

i.e. $$\sin\sigma[-\eta\zeta \sin(\alpha-\sigma) + \zeta\xi \sin(\sigma-\beta) + \xi\eta \sin(\sigma-\gamma)] \geqslant 0,$$

gives $$\xi \geqslant \frac{\eta\zeta \sin(\alpha-\sigma)}{R},$$

where $$R \equiv \eta \sin(\sigma-\gamma) + \zeta \sin(\sigma-\beta).$$

Also

$$Q - \frac{\eta\zeta \sin^2\alpha \sin(\alpha-\sigma)}{R}$$

$$\equiv \frac{1}{R}[\sqrt{\sin\sigma}\, R - \sqrt{\sin(\alpha-\sigma)\sin(\sigma-\beta)\sin(\sigma-\gamma)}\,.\,al]^2$$

$$= \frac{P^2}{R},$$

where

$$P \equiv \sqrt{\sin\sigma}\{\eta \sin(\sigma-\gamma) + \zeta \sin(\sigma-\beta)\} - \sqrt{\sin(\alpha-\sigma)\sin(\sigma-\beta)\sin(\sigma-\gamma)}\,\sqrt{\eta^2 - 2\eta\zeta\cos\alpha + \zeta^2}.$$

Further

$$\sin\sigma R^2 - \sin(\alpha-\sigma)\sin(\sigma-\beta)\sin(\sigma-\gamma)\{\eta^2 - 2\eta\zeta\cos\alpha + \zeta^2\}$$

$$= \eta^2 \sin\alpha \sin\beta \sin(\sigma-\gamma) + \zeta^2 \sin\alpha \sin\gamma \sin(\sigma-\beta) + 2\eta\zeta \sin\alpha \sin(\sigma-\beta)\sin(\sigma-\gamma)\cos(\alpha-\sigma)$$

$$\geqslant 0,$$

all the terms being positive since $\alpha - \sigma \leqslant \frac{1}{2}\pi$.

Therefore $P \geqslant 0$, and we may write the values of ξ satisfying the above two conditions in the form

(I) $$\xi \sin^2\alpha = \{\lambda P^2 + \eta\zeta \sin^2\alpha \sin(\alpha-\sigma)\}/R, \quad 0 \leqslant \lambda \leqslant 1.$$

Digitized by Google

For the value $\lambda = 0$ we have

$$D(x, y, z) = D(w, y, z) + D(w, z, x) + D(w, x, y),$$

and for the value $\lambda = 1$ we have $D(x, y, z) = 0$.

We have, further, to ensure that each of the other expressions of $D(x, y, z)$, $D(w, y, z)$, $D(w, z, x)$, $D(w, x, y)$ is not greater than the sum of the other three. We may write Δ^2 in any of the alternative forms

$$(a)\quad \Delta^2 = [\xi \sin\alpha - \eta \sin\beta - \zeta \sin\gamma]^2 - 4 \sin(\alpha - \sigma)\{-\eta\zeta \sin\sigma + \zeta\xi \sin(\sigma - \gamma) + \xi\eta \sin(\sigma - \beta)\},$$

$$(b)\quad \Delta^2 = [-\xi \sin\alpha + \eta \sin\beta - \zeta \sin\gamma]^2 + 4 \sin(\sigma - \beta)\{\eta\zeta \sin(\sigma - \gamma) - \zeta\xi \sin\sigma - \xi\eta \sin(\alpha - \sigma)\},$$

$$(c)\quad \Delta^2 = [-\xi \sin\alpha - \eta \sin\beta + \zeta \sin\gamma]^2 + 4 \sin(\sigma - \gamma)\{\eta\zeta \sin(\sigma - \beta) - \zeta\xi \sin(\alpha - \sigma) - \xi\eta \sin\sigma\}.$$

Now

$$\sin\alpha(\eta \sin\beta - \zeta \sin\gamma) - Q = \sqrt{\Omega}\sqrt{\eta^2 - 2\eta\zeta\cos\alpha + \zeta^2} - 2\sin(\alpha - \sigma)S,$$

where $$S \equiv \eta \sin(\sigma - \beta) + \zeta \sin(\sigma - \gamma),$$

and therefore is $\geqslant 0$ if and only if

$$\sin\sigma \sin(\sigma - \beta) \sin(\sigma - \gamma)\{\eta^2 - 2\eta\zeta\cos\alpha + \zeta^2\} \geqslant \sin(\alpha - \sigma)S^2,$$

i.e. if

$$\sin\alpha \sin\beta \sin(\sigma - \beta)\eta^2 + \sin\alpha \sin\gamma \sin(\sigma - \gamma)\zeta^2 \geqslant 2\eta\zeta \sin(\sigma - \beta) \sin(\sigma - \gamma)\{\sin(\alpha - \sigma) + \sin\sigma\cos\alpha\},$$

i.e.

$$\eta^2 \sin\beta \sin(\sigma - \beta) - 2\eta\zeta \sin(\sigma - \beta) \sin(\sigma - \gamma)\cos\sigma + \zeta^2 \sin\gamma \sin(\sigma - \gamma) \geqslant 0.$$

It can easily be verified that the expression on the left is not essentially positive and hence that this condition is not necessarily satisfied. If it is, then for all values of ξ given by (I),

$$\xi \sin\alpha \leqslant \eta \sin\beta + \zeta \sin\gamma,$$

and therefore

$$D(w, y, z) \not> D(x, y, z) + D(w, z, x) + D(w, x, y).$$

If it is not satisfied, then for the required result to hold we must have, by (a),

$$\xi \leqslant \frac{\eta\zeta \sin\sigma}{S},$$

whenever $$\xi \sin\alpha > \eta \sin\beta + \zeta \sin\gamma.$$

Digitized by Google

But

$$\eta \sin \beta + \zeta \sin \gamma - \frac{\eta\zeta \sin \sigma \sin \alpha}{S}$$

$$= \frac{1}{S}\{\eta^2 \sin \beta \sin (\sigma - \beta) + \zeta^2 \sin \gamma \sin (\sigma - \gamma) - 2\eta\zeta \sin (\sigma - \beta) \sin (\sigma - \gamma) \cos \sigma\}$$

$$< 0.$$

Hence in order that

$$D(w, y, z) \not> D(x, y, z) + D(w, z, x) + D(w, x, y),$$

we require

(II) If

$$\eta^2 \sin \beta \sin (\sigma - \beta) - 2\eta\zeta \sin (\sigma - \beta) \sin (\sigma - \gamma) \cos \sigma + \zeta^2 \sin \gamma \sin (\sigma - \gamma) < 0,$$

then $$\frac{\eta\zeta \sin (\alpha - \sigma)}{R} \leqslant \xi \leqslant \frac{\eta\zeta \sin \sigma}{S},$$

and it can easily be verified that this condition is stricter than condition (I).

In a similar manner we consider the condition

$$D(w, z, x) \not> D(w, y, z) + D(w, x, y) + D(x, y, z).$$

By (*b*) this requires that, if

$$\xi \sin \alpha < \eta \sin \beta - \zeta \sin \gamma,$$

we should also have

$$\eta\zeta \sin (\sigma - \gamma) - \zeta\xi \sin \sigma - \xi\eta \sin (\alpha - \sigma) \geqslant 0,$$

i.e. $$\xi \leqslant \frac{\eta\zeta \sin (\sigma - \gamma)}{\eta \sin (\alpha - \sigma) + \zeta \sin \sigma}.$$

This is certainly the case if

$$\eta \sin \beta - \zeta \sin \gamma \leqslant \frac{\eta\zeta \sin \alpha \sin (\sigma - \gamma)}{\eta \sin (\alpha - \sigma) + \zeta \sin \sigma},$$

i.e. if

$$\eta^2 \sin \beta \sin (\alpha - \sigma) - 2\eta\zeta \sin \sigma \sin (\alpha - \sigma) \cos (\sigma - \gamma) - \zeta^2 \sin \sigma \sin \gamma \leqslant 0.$$

In this case, condition (I) is sufficient. If not, we have the further condition

(III) If

$$\eta^2 \sin \beta \sin (\alpha - \sigma) - 2\eta\zeta \sin \sigma \sin (\alpha - \sigma) \cos (\sigma - \gamma) - \zeta^2 \sin \gamma \sin \sigma > 0,$$

Digitized by Google

then $$\frac{\eta\zeta \sin(\alpha-\sigma)}{R} \leqslant \xi \leqslant \frac{\eta\zeta \sin(\sigma-\gamma)}{\eta \sin(\alpha-\sigma)+\zeta \sin\sigma}.$$

Finally by considering the condition

$$D(w, x, y) \nless D(w, y, z) + D(w, z, x) + D(x, y, z),$$

we obtain

(IV) If

$$-\eta^2 \sin\beta \sin\sigma - 2\eta\zeta \sin\sigma \sin(\alpha-\sigma)\cos(\sigma-\beta) + \zeta^2 \sin\gamma \sin(\alpha-\sigma) > 0,$$

then $$\frac{\eta\zeta \sin(\alpha-\sigma)}{R} \leqslant \xi \leqslant \frac{\eta\zeta \sin(\sigma-\beta)}{\eta \sin\sigma+\zeta \sin(\alpha-\sigma)}.$$

We have finally to consider the case $\alpha=\pi$; $D(w, y, z)=0$.

$$\Delta^2 = \eta^2 \sin^2\beta + \zeta^2 \sin^2\gamma + 2\eta\zeta(1+\cos\beta\cos\gamma) - 2\xi(\eta+\zeta)(\cos\beta+\cos\gamma),$$

$\beta+\gamma<\alpha$, and therefore $\frac{1}{2}(\beta+\gamma)<\frac{1}{2}\pi$, $\frac{1}{2}|\beta-\gamma|<\frac{1}{2}\pi$.

Therefore $\cos\beta+\cos\gamma = 2\cos\frac{1}{2}(\beta+\gamma)\cos\frac{1}{2}(\beta-\gamma) \geqslant 0$,

$$\Delta^2 = [\eta\sin\beta+\zeta\sin\gamma]^2 + 2\eta\zeta(1+\cos(\beta+\gamma)) - 2\xi(\eta+\zeta)(\cos\beta+\cos\gamma)$$

$$= [\eta\sin\beta-\zeta\sin\gamma]^2 + 2\eta\zeta(1+\cos(\beta-\gamma)) - 2\xi(\eta+\zeta)(\cos\beta+\cos\gamma).$$

Proceeding as above, we find that in this case all our requirements are satisfied if

(V) $$\frac{\eta\zeta\cos\frac{1}{2}(\beta+\gamma)}{(\eta+\zeta)\cos\frac{1}{2}(\beta-\gamma)} \leqslant \xi \leqslant \frac{\eta\zeta\cos\frac{1}{2}(\beta-\gamma)}{(\eta+\zeta)\cos\frac{1}{2}(\beta+\gamma)}.$$

Putting in the values of ξ, η, ζ in terms of l, m, n, we obtain our results in the form:

(I) $$\frac{m\sin(\sigma-\beta)+n\sin(\sigma-\gamma)}{\sin(\alpha-\sigma)} \geqslant l \geqslant \frac{mn\sin^2\alpha}{m(\sin\sigma\sin(\sigma-\beta)+\sin(\alpha-\sigma)\sin(\sigma-\gamma))+n(\sin\sigma\sin(\sigma-\gamma)+\sin(\alpha-\sigma)\sin(\sigma-\beta))-\alpha\sqrt{\Omega}}.$$

Digitized by Google

(II) If

$$m^2 \sin\gamma \sin(\sigma-\gamma) - 2mn \sin(\sigma-\beta)\sin(\sigma-\gamma)\cos\sigma + n^2 \sin\beta \sin(\sigma-\beta) < 0,$$

then also $$l \geqslant \frac{m\sin(\sigma-\gamma) + n\sin(\sigma-\beta)}{\sin\sigma}.$$

(III) If

$$m^2 \sin\gamma \sin\sigma + 2mn \sin\sigma \sin(\alpha-\sigma)\cos(\sigma-\gamma) - n^2 \sin\beta \sin(\alpha-\sigma) < 0,$$

then also $$l \geqslant \frac{m\sin\sigma + n\sin(\alpha-\sigma)}{\sin(\sigma-\gamma)}.$$

(IV) If

$$-m^2 \sin\gamma \sin(\alpha-\sigma) + 2mn \sin\sigma \sin(\alpha-\sigma)\cos(\sigma-\beta) + n^2 \sin\beta \sin\sigma < 0,$$

then also $$l \geqslant \frac{m\sin(\alpha-\sigma) + n\sin\sigma}{\sin(\sigma-\beta)}.$$

(V) If $$\alpha = \pi,$$

then $$(m+n)\frac{\cos\frac{1}{2}(\beta-\gamma)}{\cos\frac{1}{2}(\beta+\gamma)} \geqslant l \geqslant (m+n)\frac{\cos\frac{1}{2}(\beta+\gamma)}{\cos\frac{1}{2}(\beta-\gamma)},$$

and $$m+n=a.$$

The two other cases $\beta > \sigma$, $\gamma > \sigma$, lead to similar results.

§ 8. *Examples illustrating the results of* § 7.

(1) Let $$\delta(w,x) = \delta(w,y) = \delta(w,z) = 1,$$
$$\delta(y,z) = \delta(z,x) = \delta(x,y) = 2.$$

Then $$\alpha = \beta = \gamma = \pi, \quad \sigma = \tfrac{3}{2}\pi.$$

The point w is common to the stretches joining the points yz, zx, xy.

$$D(w,y,z) = D(w,z,x) = D(w,x,y) = 0,$$
$$D(x,y,z) = \sqrt{3}.$$

(2) If we take $$\alpha = \frac{\pi}{2}, \quad \beta = \gamma = \frac{\pi}{6},$$

then $$\sigma = \frac{5\pi}{12}, \quad \alpha - \sigma = \frac{\pi}{12}, \quad \sigma - \beta = \sigma - \gamma = \frac{\pi}{4}.$$

Digitized by Google

The conditions to be satisfied by ξ for the case $\eta = \zeta = 1$ are those numbered (I) and we must have

$$\frac{\sqrt{3}-1}{4} \leqslant \xi \leqslant \sqrt{3} - 1.$$

Thus, if we take $\xi = \frac{1}{2}$, $l = \sqrt{2}$, $m = n = \dfrac{1}{\sqrt{2}}$,

therefore $a = 1$, $b = c = \sqrt{\frac{5}{2} - \sqrt{3}} = \cdot 876\ldots.$

$$D(w, y, z) = D(w, z, x) = D(w, x, y) = \tfrac{1}{4},$$

$$D(x, y, z) = \sqrt{\tfrac{9}{4} - \sqrt{3}} = \cdot 36\ldots.$$

(3) If in the above example we take

$$\xi = \frac{\sqrt{3}-1}{4},$$

$$D(x, y, z) = D(w, y, z) + D(w, z, x) + D(w, x, y).$$

The corresponding values are

$$l = \sqrt{2 + 2\sqrt{3}} = 2 \cdot 34\ldots,$$

$$m = n = \sqrt{\frac{\sqrt{3}-1}{4}} = \cdot 43\ldots,$$

$$a = \sqrt{\frac{\sqrt{3}-1}{2}} = \cdot 605\ldots, \quad b = c = 1 \cdot 98\ldots.$$

The point w belongs to the triangle defined by x, y, z.

(4) If we take $\alpha = \pi$, $\beta = \gamma = 0$, then

$$D(w, y, z) = D(w, z, x) = D(w, x, y) = 0$$

and $$D(x, y, z) = \sqrt{lmn(m + n - l)},$$

and therefore for the space to be metrical we require

$$m + n \geqslant l, \quad a = m + n, \quad b = l - n, \quad c = l - m.$$

Hence y and z both lie on the stretch xw, while w lies on the stretch yz.

We notice that in this example, where the condition

$$D(x, y, z) \leqslant D(w, y, z) + D(w, z, x) + D(w, x, y)$$

is not satisfied, the stretches joining the points concerned are not linear.

We can indeed shew that in general whenever the condition is satisfied for any four points (whether the triangulation is plane or not) then the stretches are linear. We have already proved the

Digitized by Google

result for the case where the triangulation is plane. Now let x, y be any two points and z, w any two points of the stretch joining x and y. Thus, let

$$\delta(x, y) = l, \quad \delta(z, w) = a, \quad \delta(z, x) = b, \quad \delta(x, w) = c.$$

Then also

$$\delta(y, w) = l - c, \quad \delta(y, z) = l - b, \quad \hat{xwy} = \hat{xzy} = \pi,$$

and therefore $$D(w, x, y) = D(x, y, z) = 0.$$

Hence the requisite condition can be satisfied only if

$$D(w, y, z) = D(w, z, x).$$

$$2b^2c^2 + 2c^2a^2 + 2a^2b^2 - a^4 - b^4 - c^4$$
$$= 2(l-c)^2(l-b)^2 + 2a^2\{(l-b)^2 + (l-c)^2\} - a^4 - (l-b)^4 - (l-c)^4,$$

i.e. $$l\{a^2 - (b-c)^2\} = (b+c)\{a^2 - (b-c)^2\}.$$

In general, therefore, we must have

$$a^2 = (b-c)^2,$$

i.e. $$\delta(w, z) = |\delta(x, z) - \delta(x, w)|.$$

In the special case $b + c = l$ this result does not necessarily follow; but if the extended stretch defined by xy contains a point other than the points of the stretch joining x and y, we can deduce the required result. For let t be a point of the extended stretch not belonging to xy, and suppose it is such that x lies on ty (the proof in the case where y lies on xt is exactly similar). Then by § 1 (2) every point of xy belongs to ty and further x lies on both tw and tz. Hence if $\delta(t, x) = d > 0$,

$$\delta(t, w) = c' = c + d, \quad \delta(t, z) = b' = b + d, \quad \delta(t, y) = l' = l + d.$$

If $$b + c = l,$$

then $$b' + c' = b + c + 2d \neq l',$$

and therefore as above

$$\delta(z, w) = |\delta(t, w) - \delta(t, z)| = |\delta(x, w) - \delta(x, z)|.$$

We have now proved that for all points w, z, we have

$$\delta(z, w) = |\delta(x, z) - \delta(x, w)|,$$

and hence the stretch xy is linear.

§ 9. *Complete conditions for*

$$D(x, y, z) \leqslant D(w, y, z) + D(w, z, x) + D(w, x, y) \text{ etc.}$$

Denoting these values of the function $D(x, y, z)$ by D, D_1, D_2,

Digitized by Google

D_3, the necessary and sufficient condition that each should be less than or equal to the sum of the other three is

$$(D_1 + D_2 + D_3 - D)(D_1 + D_2 - D_3 + D) \times (D_1 - D_2 + D_3 + D)(-D_1 + D_2 + D_3 + D) \geqslant 0,$$

i.e.
$$-\Sigma D^4 + 2\Sigma D_r^2 D_s^2 + 8DD_1D_2D_3 \geqslant 0.$$

Putting in the values of D, D_1, D_2, D_3 in terms of $\xi, \eta, \zeta, \alpha, \beta, \gamma$. we have

$$-[\Sigma\xi^2 \sin^2\alpha + 2\Sigma\eta\zeta(\cos\beta\cos\gamma - \cos\alpha)]^2 - \Sigma\xi^4 \sin^4\alpha + 2\Sigma\eta^2\zeta^2 \sin^2\beta \sin^2\gamma + 2[\Sigma\xi^2 \sin^2\alpha][\Sigma\xi^2 \sin^2\alpha + 2\Sigma\eta\zeta(\cos\beta\cos\gamma - \cos\alpha)] + 128DD_1D_2D_3 \geqslant 0,$$

i.e.
$$4(1 - \cos^2\alpha - \cos^2\beta - \cos^2\gamma + 2\cos\alpha\cos\beta\cos\gamma)\Sigma\eta^2\zeta^2 + 128DD_1D_2D_3 \geqslant 0,$$

and, putting in the values of ξ, η, ζ in terms of l, m, n, this becomes

$$D(x, y, z)\, D(w, y, z)\, D(w, z, x)\, D(w, x, y) + \tfrac{9}{8}(l^2 + m^2 + n^2)[V(x, y, z, w)]^2 \geqslant 0.$$

Hence we have the two sets of conditions:

(1) $$[V(x, y, z, w)]^2 \geqslant 0,$$

in which case the triangulation is plane;

(2) $$[V(x, y, z, w)]^2 < 0$$

and

$$D(x, y, z)\, D(w, y, z)\, D(w, z, x)\, D(w, x, y) \geqslant \tfrac{9}{8}(l^2 + m^2 + n^2)[-V^2],$$

in which case the triangulation is not plane.

The two cases correspond, roughly speaking, to the cases of developable and non-developable surfaces in euclidean geometry. If we take as our space Σ a set of points forming a two-dimensional surface in euclidean space we can, under certain conditions, take as our definition of δ the geodesic distance between the points. If the surface is developable, it can be applied to a euclidean plane; this corresponds to case (1). If the surface is not developable we should not expect the function D to have regular properties, but case (2) seems to indicate that there may exist such surfaces for which D has at least the properties of § 4.

Digitized by Google

Note on Ramanujan's arithmetical function $\tau(n)$. By Mr G. H. HARDY, Trinity College.

[*Received* 4 January, *read* 31 January 1927.]

1. In his remarkable memoir 'On certain arithmetical functions'* Ramanujan considers, among other functions of much interest, the function $\tau(n)$ defined by

$$f(x) = x\{(1-x)(1-x^2)(1-x^3)\dots\}^{24} = \Sigma\tau(n)\,x^n \quad \dots(1{\cdot}1).$$

This function is important in the theory of the representation of a number as a sum of 24 squares. In fact

$$(1+2x+2x^4+2x^9+\dots)^{24} = \Sigma r_{24}(n)\,x^n = \Sigma\delta_{24}(n)\,x^n + \Sigma e_{24}(n)\,x^n:$$

where $r_{24}(n)$ is the number of representations;

$$\tfrac{691}{16}\,\delta_{24}(n) = \sigma_{11}(n) - 2\sigma_{11}'(\tfrac{1}{2}n)\dagger,$$

where $\sigma_s(n)$ is the sum of the sth powers of the divisors of n, and $\sigma_s'(n)$ the sum of those of its odd divisors; and

$$\tfrac{691}{128}\,e_{24}(n) = (-1)^{n-1}\,259\tau(n) - 512\tau(\tfrac{1}{2}n)\dagger.$$

2. The associated Dirichlet's series

$$F(s) = \Sigma\frac{\tau(n)}{n^s} \quad \dots\dots(2{\cdot}1)$$

is convergent for sufficiently large positive s. Ramanujan arrived by conjecture at the very remarkable identity

$$F(s) = \Pi\left\{\frac{1}{1-\tau(p)\,p^{-s}+p^{11-2s}}\right\} \quad \dots\dots(2{\cdot}2),$$

where the product extends over all primes p; and a proof of this formula has since been given by Mordell‡. It follows that, if we write $n = p_1^{a_1}p_2^{a_2}\dots p_r^{a_r}$ and

$$\cos\theta_p = \tfrac{1}{2}p^{-\frac{1}{2}}\,\tau(p) \quad \dots\dots(2{\cdot}3),$$

then

$$\tau(n) = n^{\frac{11}{2}}\,\frac{\sin(a_1+1)\theta_{p_1}}{\sin\theta_{p_1}}\dots\frac{\sin(a_r+1)\theta_{p_r}}{\sin\theta_{p_r}} \quad \dots(2{\cdot}4),$$

and $\tau(n)\,\tau(n') = \tau(nn')$ when n and n' are coprime.

* S. Ramanujan, 'On certain arithmetical functions', *Trans. Camb. Phil. Soc.*, 22 (1916), 159–184.

† Ramanujan, *l.c.*, 179, 184. A function with argument $\frac{1}{2}n$ is zero when n is odd.

‡ L. J. Mordell, 'On Mr Ramanujan's empirical expansions of modular functions', *Proc. Camb. Phil. Soc.*, 19 (1920), 117–124 (a paper communicated in 1917).

Digitized by Google

3. The problem of determining the order of magnitude of $\tau(n)$ appears to be very difficult. Ramanujan proved, on the one hand that

$$\tau(n) = O(n^7)^* \qquad \ldots\ldots(3{\cdot}1),$$

and on the other that

$$\tau(n) = \Omega(n^5)\dagger \qquad \ldots\ldots(3{\cdot}2).$$

He also proved‡ that, if we assume the truth of (2·2), an assumption since justified by Mordell, and also of the inequality

$$\tau(p) \leqq 2p^{\frac{11}{2}} \qquad \ldots\ldots(3{\cdot}3),$$

then

$$\tau(n) \leqq n^{\frac{11}{2}} d(n) \qquad \ldots\ldots(3{\cdot}4)$$

and

$$\tau(n) \geqq n^{\frac{11}{2}} \qquad \ldots\ldots(3{\cdot}5)$$

for an infinity of values of n, so that (3·1) and (3·2) may be replaced by

$$\tau(n) = O(n^{\frac{11}{2}+\epsilon}) \qquad \ldots\ldots(3{\cdot}6),$$

for every positive ϵ, and

$$\tau(n) = \Omega(n^{\frac{11}{2}}) \qquad \ldots\ldots(3{\cdot}7).$$

If then (3·3) were proved, the problem would be in essentials solved.

4. In this note I make three contributions to the problem. The first (which is all but trivial) is to show that (3·7) does not depend upon the unproved inequality (3·3) but only on the identities (2·2) and (2·4) established by Mordell, and is therefore certainly true. The second is to show that

$$\tau(n) = O(n^6) \qquad \ldots\ldots(4{\cdot}1),$$

which therefore replaces Ramanujan's (3·1); and the third is to show that, if

$$T(n) = \{\tau(1)\}^2 + \{\tau(2)\}^2 + \ldots + \{\tau(n)\}^2 \qquad \ldots\ldots(4{\cdot}2),$$

then there are positive constants A and B such that

$$An^{12} < T(n) < Bn^{12} \qquad \ldots\ldots(4{\cdot}3)$$

for $n \geqq 1$. This shows that the average order of $\tau(n)$ is exactly $O(n^{\frac{11}{2}})$, and also gives an alternative proof of both (3·7) and (4·1), and indeed of more.

* Ramanujan, *l.c.*, 168, 171.
† That is to say $\tau(n) \neq o(n^5)$. See Ramanujan, *l.c.*, 171.
‡ Ramanujan, *l.c.*, 175.

Digitized by Google

Proof of (3·7).

5. Ramanujan shows that (3·7) follows from (2·2) and (3·3). It is therefore only necessary to show that (3·7) is true if (3·3) is false for some p.

If (3·3) is false, $\cos\theta_p$ is real and θ_p complex, so that $\theta_p = k\pi + i\eta_p$, where k is an integer and η_p is real. We may suppose η_p positive, so that $e^{\eta_p} > 1$. Then, if $n = p^a$, we have

$$|\tau(n)| = n^{\frac{11}{2}} \left|\frac{\sin(a+1)\theta_p}{\sin\theta_p}\right| = n^{\frac{11}{2}} \left|\frac{\sinh(a+1)\eta_p}{\sinh\eta_p}\right|$$

$$> An^{\frac{11}{2}} e^{a\eta_p} = An^{\frac{11}{2}+\delta},$$

where A is a positive constant and $\delta = \eta_p/\log p > 0$, for all values of a; and this evidently completes the proof of (3·7).

Proof of (4·1).

6·1. The proof of (4·1) depends upon the following lemma, which is interesting in itself.

Lemma 1. *If* $f(x)$ *is the function* (1·1), *then*

$$f(x) = f(re^{i\theta}) = O\{(1-r)^{-6}\} \qquad \ldots\ldots(6{\cdot}11)$$

uniformly in θ.

The proof of this lemma depends on the methods used by Littlewood and myself in one of our memoirs on Diophantine approximation*. Following the notations of that memoir, I write

$$x = q^2, \quad f(x) = \phi(q), \quad q = e^{\pi i\tau}, \quad \mathrm{T} = \frac{c+d\tau}{a+b\tau}, \quad Q = e^{\pi i\mathrm{T}},$$

where a, b, c, d are integers such that $ad - bc = 1$. Then the equation of transformation for $\phi(q)$ is

$$\phi(Q) = (a+b\tau)^{12}\phi(q).$$

Hence, if

$$\chi(q) = \mathfrak{R}\log\phi(q),$$

we have

$$\chi(q) = \chi(Q) - 12\log|a+b\tau|.$$

If we continue to follow both the ideas and the notation of the memoir referred to†, in which in particular

$$e^{-\pi y} = r, \quad \pi y = \log\frac{1}{r} \sim 1-r, \quad |Q| = e^{-\pi\lambda},$$

* G. H. Hardy and J. E. Littlewood, 'Some problems of Diophantine approximation: II: The trigonometrical series associated with the elliptic theta-functions', *Acta Math.*, 37 (1914), 193–238.

† 226 *et seq.*

Digitized by Google

we obtain

$$\chi(q) = \chi(Q) - 6\log\left(\frac{1}{q'^2_{n+1}} + q_n^2 y^2\right) = \chi(Q) - 6\log y - 6\log\frac{1}{\lambda},$$

$$\chi(q) + 6\log y = -6\log(1/\lambda) + \Re\log\{Q^2(1-Q^2)^{24}(1-Q^4)^{24}\ldots\}$$

$$= -6\log\frac{1}{\lambda} - 2\pi\lambda + 24\,\Sigma\Re\log(1-Q^{2n}) < A - 6\log\frac{1}{\lambda} - 2\pi\lambda < A,$$

where the A's are constants*. But this is equivalent to

$$|f(x)| = |\phi(q)| < Ay^{-6} < A\left(\log\frac{1}{r}\right)^{-6} < A(1-r)^{-6},$$

which proves the lemma.

6·2. We can deduce (4·1) at once from (6·11). For

$$\tau(n) = \frac{1}{2\pi i}\int\frac{f(x)}{x^{n+1}}\,dx = O\{(1-r)^{-6}\} = O(n^6),$$

if the integration is effected round the circle of radius $r = 1 - 1/n$.

Proof of (4·3).

7·1. To prove (4·3) we require a further lemma.

Lemma 2. *If* $a_n \geqq 0$, $h > 0$, $k > 0$, $\alpha > 0$ *and*

$$h(1-r)^{-\alpha} < g(r) = \Sigma a_n r^n < k(1-r)^{-\alpha} \quad \ldots\ldots(7{\cdot}11)$$

for all values of r *less than and sufficiently near to* 1, *then there are positive constants* p *and* q *such that*

$$pn^\alpha < s_n = a_0 + a_1 + a_2 + \ldots + a_n < qn^\alpha \quad \ldots\ldots(7{\cdot}12)$$

for all sufficiently large values of n.

The second inequality (7·12) is immediate, since

$$s_n \leqq r^{-n}\sum_0^n a_\nu r^\nu \leqq r^{-n} g(r) \leqq 4g\left(1 - \frac{1}{n}\right) < 4kn^\alpha$$

if $r = 1 - 1/n$ and n is sufficiently large. The first inequality is not quite so obvious.

We write $r = e^{-y}$, so that $1 - r \sim y$. Then plainly

$$G(y) = \Sigma s_n e^{-ny} > \tfrac{1}{2}hy^{-\alpha-1} \quad \ldots\ldots(7{\cdot}13)$$

if y is sufficiently small. We choose c so that

$$c > 2\alpha, \quad 8qc^\alpha e^{-c} < h \quad \ldots\ldots(7{\cdot}14),$$

* The function $-6\log(1/\lambda) - 2\pi\lambda$ becomes negatively infinite when $\lambda \to 0$ or $\lambda \to \infty$, and has a maximum when $\pi\lambda = 3$. Here $\lambda > \frac{1}{2}$.

Digitized by Google

and take $y = c/\nu$, where ν is a large integer; and we write

$$G(y) = \sum_1^{\nu} s_n e^{-ny} + \sum_{\nu+1}^{\infty} s_n e^{-ny} = G_1(y) + G_2(y) \quad \ldots(7{\cdot}15).$$

Then

$$G_2(y) < q \sum_{\nu+1}^{\infty} n^\alpha e^{-ny} < q \int_\nu^\infty u^\alpha e^{-uy}\, du = qy^{-\alpha-1} \int_c^\infty w^\alpha e^{-w}\, dw,$$

if ν is sufficiently large *. But

$$\chi(c) = \int_c^\infty w^\alpha e^{-w}\, dw = c^\alpha e^{-c} + \alpha \int_c^\infty w^{\alpha-1} e^{-w}\, dw < c^\alpha e^{-c} + \tfrac{1}{2}\chi(c),$$

since $c > 2\alpha$; and so

$$G_2(y) < 2qc^\alpha e^{-c} y^{-\alpha-1} < \tfrac{1}{4} h y^{-\alpha-1} \quad \ldots\ldots(7{\cdot}16),$$

by (7·14). From (7·13)—(7·16) we deduce

$$G_1(y) = G(y) - G_2(y) > \tfrac{1}{4} h y^{-\alpha-1} \quad \ldots\ldots(7{\cdot}17).$$

But

$$G_1(y) = \sum_{n \leqq \nu} s_n e^{-ny} \leqq \frac{s_\nu}{1 - e^{-y}} \leqq \frac{2 s_\nu}{y} \quad \ldots\ldots(7{\cdot}18),$$

if y is small enough. Comparing (7·17) and (7·18) we see that

$$s_\nu > \tfrac{1}{8} h y^{-\alpha} = \tfrac{1}{8} h c^{-\alpha} \nu^\alpha$$

for sufficiently large ν, which proves the first inequality (7·11).

7·2. The second inequality (4·3) is an immediate deduction from Lemmas 1 and 2. For

$$g(r) = \Sigma \{\tau(n)\}^2 r^{2n} = \frac{1}{2\pi} \int_0^{2\pi} |f(re^{i\theta})|^2\, d\theta$$

$$= O\{(1-r)^{-12}\} = O(n^{12})$$

when $r = 1 - 1/n$, by Lemma 1; and $\mathrm{T}(n) < Bn^{12}$ by Lemma 2. This inequality plainly includes (4·1), since $\mathrm{T}(n) \geqq \{\tau(n)\}^2$.

To prove the first inequality (4·3) we use Jacobi's identity †

$$\psi(x) = \{(1-x)(1-x^2)(1-x^3)\ldots\}^3 = 1 - 3x + 5x^3 - 7x^6 + \ldots,$$

where the indices are the triangular numbers.

Since $f = x\psi^8$, we have

$$g(r) = \frac{r^2}{24} \int_0^{2\pi} |\psi(re^{i\theta})|^{16}\, d\theta > A \left(\int_0^{2\pi} |\psi(re^{i\theta})|^2\, d\theta \right)^8$$

$$> A(1 + 3^2 r^2 + 5^2 r^6 + 7^2 r^{12} + \ldots)^8$$

* The maximum of $u^\alpha e^{-uy}$ occurs for $u = \alpha/y$, outside the range of integration because $c > \alpha$.

† *Fundamenta nova*, § 66.

Digitized by Google

when $r \to 1$. But if $r = e^{-\delta}$, we have

$$\Sigma (2n+1)^2 r^{n(n+1)} \sim \tfrac{1}{4} \Sigma n^2 r^{n^2} \sim \tfrac{1}{4} \int_0^\infty t^2 e^{-\delta t^2} dt$$

$$\sim A\delta^{-\frac{3}{2}} \sim A (1-r)^{-\frac{3}{2}},$$

and so

$$g(r) > A (1-r)^{-12};$$

and therefore, by Lemma 2, $\mathrm{T}(n) > An^{12}$.

Since

$$\mathrm{T}(n) \leqq n \operatorname*{Max}_{\nu \leqq n} |\tau(\nu)|^2,$$

this gives an alternative proof of (3·7).

[Added 9 Feb. 1927.] The note was written nearly 10 years ago, my interest in the matter reviving recently as a result of editorial work in connection with the forthcoming edition of Ramanujan's works. The index 6 of (4·1) is the $\frac{1}{4}r$ which occurs in the recent work of Landau, Petersson, and Walfisz concerning the number of representations of n by r squares, and the associated lattice-point problem; and (4·1) itself must be included implicitly in their general results. See H. Petersson, 'Über die Anzahl der Gitterpunkte in mehrdimensionalen Ellipsoiden', *Hamburg Math. Abhandlungen*, 5 (1926), 116–150; and the memoirs of Landau and Walfisz there referred to.

Digitized by Google

Notes on the theory of series (III): *On the summability of the Fourier series of a nearly continuous function.* By Mr G. H. HARDY and Mr J. E. LITTLEWOOD.

[*Received* 4 January, *read* 31 January 1927.]

1. The theorem which we prove here seems obvious enough when stated, but it appears to have been overlooked by the numerous writers who have discussed the subject, and the proof is less immediate than might be expected.

We say that $\phi(t)$ is continuous (C, α), where $\alpha > 0$, if

$$\frac{1}{\Gamma(\alpha)\, t^{\alpha}} \int_0^t \phi(u)\,(t-u)^{\alpha-1}\, du$$

tends to a finite limit l when $t \to 0$. We shall also say that $\phi(t) \to l\ (C, \alpha)$. To assert continuity (C, α), for a small α, is to assert a little less than ordinary continuity.

If $f(t)$ is periodic and integrable, and

$$\phi(t) = \tfrac{1}{2}\{f(x+t) + f(x-t) - 2s\} \to 0,$$

then the Fourier series of $f(t)$, for $t = x$, is summable (C, δ) to sum s for every positive δ.* If $\phi(t) \to 0\ (C, 1)$, then the series is summable $(C, 1+\delta)$.† It is therefore natural to expect the truth of the following theorem.

THEOREM. *If* $\phi(t) \to 0\ (C, \alpha)$, *then the Fourier series of* $f(t)$, *for* $t = x$, *is summable* $(C, \alpha + \delta)$ *to sum* s *for every positive* δ.

We prove the theorem in the most interesting case when $0 < \alpha < 1$. The proof for $\alpha > 1$ will differ in complication but not in principle.

2. We suppose β and δ positive and $\beta + \delta < 1$. The $(\beta+\delta)$-th Rieszian mean of the Fourier series, for terms whose rank does not exceed ω, differs from s by

$$\frac{2\Gamma(1+\beta+\delta)}{\pi} J_{\beta+\delta}(\phi, \omega) = \frac{2\Gamma(1+\beta+\delta)}{\pi} \int_0^{\infty} t^{-1-\beta-\delta}\, C_{1+\beta+\delta}(t)\, \phi\left(\frac{t}{\omega}\right) dt,$$

where $C_q(t)$ is Young's function‡

$$C_q(t) = \frac{t^q}{\Gamma(1+q)} \left\{1 - \frac{t^2}{(1+q)(2+q)} + \frac{t^4}{(1+q)\dots(4+q)} - \dots\right\}.$$

Since $$C_{1+\beta+\delta}(\omega u) = \frac{\omega^{\beta}}{\Gamma(\beta)} \int_0^u C_{1+\delta}(\omega v)\,(u-v)^{\beta-1} dv,$$

* M. Riesz, **4**, **5**; Chapman, **1**. See also Hobson, **3**.
† Young, **6**, **8**. ‡ Young, **7**.

Digitized by Google

we have

$$J_{\beta+\delta}(\phi, \omega) = \omega^{-\beta-\delta}\int_0^\infty u^{-1-\beta-\delta}C_{1+\beta+\delta}(\omega u)\,\phi(u)\,du$$

$$= \frac{\omega^{-\delta}}{\Gamma(\beta)}\int_0^\infty u^{-1-\beta-\delta}\phi(u)\,du\int_0^u C_{1+\delta}(\omega v)(u-v)^{\beta-1}dv$$

$$= \frac{\omega^{-\delta}}{\Gamma(\beta)}\int_0^\infty C_{1+\delta}(\omega v)\,dv\int_v^\infty u^{-1-\beta-\delta}(u-v)^{\beta-1}\phi(u)\,du$$

$$= \frac{\omega^{-\delta}}{\Gamma(\beta)}\int_0^\infty v^{-1-\delta}C_{1+\delta}(\omega v)\,\psi(v)\,dv = \frac{1}{\Gamma(\beta)}J_\delta(\psi, \omega),$$

where $$\psi(v) = v^{1+\delta}\int_v^\infty u^{-1-\beta-\delta}(u-v)^{\beta-1}\phi(u)\,du.$$

There is no difficulty in the inversion of the order of integration, since $C_{1+\delta}(t)$ is bounded, and $O(t^{1+\delta})$ for small t, and the double integral is absolutely convergent.

Assume for the moment that it has been proved that $\psi(v)$ tends to zero with v. Then $J_\delta(\psi, \omega)\to 0$ when $\omega\to\infty$, this being indeed the kernel of the proof that the Fourier series of a continuous function is summable (C, δ).* Hence $J_{\beta+\delta}(\phi, \omega)\to 0$, and the Fourier series of $f(t)$ is summable $(C, \beta+\delta)$ for $t = x$. The theorem will accordingly be proved if we can shew that $\psi(v)\to 0$ with v whenever $\beta > \alpha$ and $\delta > 0$.

3. We may suppose without loss of generality that $x = 0$, that $s = 0$, and that the Fourier series of $f(t)$ has no constant term, so that the mean value of $\phi(t)$ is zero. It is in fact easily verified that $\psi(v)\to 0$ with v when $f(t)$ is $f_0(t) = a_0 - (a_0 - s)\cos t$, and we may consider $f - f_0$ instead of f. Our hypothesis is that

$$\phi_\alpha(t) = A\int_0^t \phi(u)(t-u)^{\alpha-1}du = o(t^\alpha), \quad \ldots\ldots(3{\cdot}1)$$

A being, here and in the sequel, a constant (a different constant in different places). From (3·1) it follows that

$$\phi(t) = A\frac{d}{dt}\int_0^t \frac{\phi_\alpha(u)}{(t-u)^\alpha}du = \phi_1'(t), \quad \ldots\ldots\ldots(3{\cdot}2)$$

* See Young, **6**, **7**; Hardy, **2**. Here ψ will play the part of the ϕ of the ordinary proof. The genesis of ψ is really irrelevant, and it is not important that ψ is not periodic; but we can, if we please, replace ψ by the periodic function ψ^* equal to ψ in $(0, 2\pi)$, it being easily proved that the difference between $J_\delta(\psi, \omega)$ and $J_\delta(\psi^*, \omega)$, for any ψ integrable in $(0, 2\pi)$, is $O(\omega^{-\delta})$ and therefore trivial.

Digitized by Google

say, for almost all t.* Also

$$\phi_1(t) = \int_0^t \frac{o(u^\alpha)}{(t-u)^\alpha}\,du = o(t) \quad \ldots\ldots\ldots\ldots\ldots\ldots (3{\cdot}3)$$

for small t, while, being the integral of $\phi(t)$, it is $O(1)$ for large t.

We write

$$\psi(v) = v^{1+\delta}\int_v^\infty = v^{1+\delta}\int_v^{2v} + v^{1+\delta}\int_{2v}^\infty = \psi_1 + \psi_2,$$

and we dispose first of ψ_2. This is

$$-v^{1+\delta}(2v)^{-1-\beta-\delta}v^{\beta-1}\phi_1(2v) - v^{1+\delta}\int_{2v}^\infty \phi_1(u)\frac{d}{du}\{u^{-1-\beta-\delta}(u-v)^{\beta-1}\}\,du.$$

The first term is $o(1)$, by (3·3). The second is

$$\begin{aligned} &v^{1+\delta}\int_{2v}^1 o(u)\,O\{u^{-2-\beta-\delta}(u-v)^{\beta-1}\}\,du \\ &\qquad + v^{1+\delta}\int_{2v}^1 o(u)\,O\{u^{-1-\beta-\delta}(u-v)^{\beta-2}\}\,du \\ &\qquad + v^{1+\delta}\int_1^\infty O(1)\,O(u^{-3-\delta})\,du \\ &= o(1) + o(1) + O(v^{1+\delta}) = o(1). \end{aligned}$$

Hence $$\psi_2 \to 0.$$

4. It remains to prove that $\psi_1 \to 0$. We observe first that if $0 < \xi \leqq \eta \leqq 2\xi$, then

$$\int_\xi^\eta \phi(u)\,du = o\{\xi^\alpha(\eta-\xi)^{1-\alpha}\}.$$

We have in fact

$$\phi_1(\eta) - \phi_1(\xi) = A\int_\xi^\eta \frac{\phi_\alpha(t)}{(\eta-t)^\alpha}\,dt - A\int_0^\xi \phi_\alpha(t)\left\{\frac{1}{(\xi-t)^\alpha} - \frac{1}{(\eta-t)^\alpha}\right\}dt.$$

The first term here is

$$o\left\{\xi^\alpha\int_\xi^\eta \frac{dt}{(\eta-t)^\alpha}\right\} = o\{\xi^\alpha(\eta-\xi)^{1-\alpha}\};$$

and the second is

$$o\left[\xi^\alpha\int_0^\xi\left\{\frac{1}{(\xi-t)^\alpha} - \frac{1}{(\eta-t)^\alpha}\right\}dt\right] = o[\xi^\alpha\{(\eta-\xi)^{1-\alpha} - \eta^{1-\alpha} + \xi^{1-\alpha}\}],$$

which is of the same form; and this establishes our assertion.

* This is simply the solution of 'Abel's integral equation'.

Digitized by Google

We have now

$$\int_v^{2v} u^{-1-\beta-\delta}(u-v)^{\beta-1}\phi(u)\,du = \left[u^{-1-\beta-\delta}(u-v)^{\beta-1}\left(\int_v^u \phi(t)\,dt\right)\right]_v^{2v}$$

$$-\int_v^{2v}\left[\frac{d}{du}\left\{u^{-1-\beta-\delta}(u-v)^{\beta-1}\right\}\int_v^u \phi(t)\,dt\right]du$$

$$= o\left(v^{\alpha}.v^{1-\alpha}.v^{-1-\delta-\delta}.v^{\beta-1}\right)$$

$$+\int_v^{2v} O\left\{u^{-2-\beta-\delta}(u-v)^{\beta-1}+u^{-1-\beta-\delta}(u-v)^{\beta-2}\right\} o\left\{v^{\alpha}(u-v)^{1-\alpha}\right\}du$$

$$= \int_v^{2v} o\left\{v^{-2-\beta-\delta+\alpha}(u-v)^{\beta-\alpha}+v^{-1-\beta-\delta+\alpha}(u-v)^{\beta-\alpha-1}\right\}du$$

$$= o\left(v^{-2-\beta-\delta+\alpha}.v^{\beta-\alpha+1}+v^{-1-\beta-\delta+\alpha}.v^{\beta-\alpha}\right) = o\left(v^{-1-\delta}\right),$$

or $\psi_2 = o(1)$; which completes the proof.

REFERENCES

1. S. Chapman, 'On non-integral orders of summability of series and integrals', *Proc. London Math. Soc.* (2), 9 (1910), 369–409.
2. G. H. Hardy, 'On the summability of Fourier series', *Proc. London Math. Soc.* (2), 12 (1913), 365–372.
3. E. W. Hobson, *The theory of functions of a real variable*, 2 (ed. 2, 1926), 567.
4. M. Riesz, 'Sur les séries de Dirichlet et les séries entières', *Comptes rendus*, 22 Nov. 1909.
5. M. Riesz, 'Sur la sommation des séries de Fourier', *Acta Reg. Univ. Francisco-Josephinae* (Szeged), 1 (1923), 104–113.
6. W. H. Young, 'Über eine Summationsmethode für die Fouriersche Reihe', *Leipziger Berichte*, 63 (1911), 369–387.
7. W. H. Young, 'On infinite integrals involving a generalisation of the sine and cosine functions', *Quarterly Journal*, 43 (1912), 161–177.
8. W. H. Young, 'On the convergence of a Fourier series and of its allied series', *Proc. London Math. Soc.* (2), 10 (1912), 254–272.

Digitized by Google

The evaluation of Gibbs' phase-integral for imperfect gases. By H. D. URSELL, Trinity College. (Communicated by Mr R. H. FOWLER.)

[*Read* 31 January, *received* 1 February 1927.]

Introduction.

Statistical mechanics is concerned primarily with what are known as "normal properties" of assemblies. The underlying idea is that of the generalised phase-space. The configuration of an assembly is determined (on classical mechanics) by a certain number of pairs of Hamiltonian canonical coordinates p, q, which are the coordinates of the phase-space referred to. Liouville's theorem leads us to take the element of volume $d\tau = \Pi\, dp\, dq$ as giving the correct element of *a priori* probability. Any isolated assembly is confined to a surface in the phase-space, for its energy at least is constant; when there are no other uniform integrals of the equations of motion, the actual probability of a given aggregate of states of the proper energy, i.e., of a given portion of the surface, varies as the volume, in the neighbourhood of points of this portion, included between two neighbouring surfaces of constant energies E, $E + dE$; it therefore varies as the integral of $(\partial E/\partial n)^{-1}$ taken over the portion. If I be the measure of the total phase-space available, interpreted in this way, and i that of the portion in which some particular condition is satisfied, then i/I is the probability of that condition being satisfied.

The importance of the idea lies in the fact that we can observe experimentally only a few functions of the canonical variables, and all points of the phase-space for which these have the same values are indistinguishable to us. We may therefore take any set of such measurable quantities as coordinates in a second space, and we can then calculate the probability of any given region in this space. A particular set of values of the measurables is said to be a normal property of the assembly if i/I has effectively the value unity for a region of negligible dimensions surrounding the point giving these values in the second space. Clearly a normal property must necessarily be an average property; moreover properties are only measurable when they are normal.

The present paper arose in a recalculation of such a ratio in a case in which the usual method seems unjustifiable. On the basis of Maxwell's law of velocity-distribution we can calculate the pressure of an ideal gas composed of point molecules, obtaining the usual law

$$p = nkT,$$

Digitized by Google

where n is the molecular density. In applying the same method to any other molecular model we are led at once into complications. Taking the usual picture of hard spherical molecules, of diameter D, and a hard elastic wall we find that the old formula applies provided that n denotes the molecular density at the boundary (strictly at a distance $\frac{1}{2}D$ from the boundary). Here n is the density at any particular point if $nIdv$ N is the measure i of the available phase-space for which an assigned molecule has its centre in a small volume dv surrounding the point. This will not be the same when the point is at a distance $\frac{1}{2}D$ from the boundary as when it is in the interior of the gas.

The simplest way of obtaining a first approximation is perhaps as follows*. In the first place we may ignore the impulse co-ordinates, which give a factor in i or I quite independent of the other coordinates. Let V be the volume available for molecular centres; then by imagining the N molecules placed in position in the volume V in succession we find

$$I = V(V-a)(V-2a)\ldots\ldots\{V-(N-1)a\},$$

where a is eight times the volume of a molecule. Writing V' for $V-\frac{1}{2}a$ we find in the same way

$$i = dv\,.\,V'(V'-a)(V'-2a)\ldots\ldots\{V'-(N-2)a\},$$

when dv is at the boundary. If it were in the interior we should have to put $V-a$ for V'. Thus

$$\frac{nV}{N} = \frac{iV}{Idv} = \left(\frac{V'}{V-a}\right)^N \prod_{r=1}^{N-1} \frac{(V+\frac{1}{2}a-ra)(V-a)}{(V-ra)(V-\frac{1}{2}a)} = 1+\frac{b}{V},$$

where $b = \frac{1}{2}Na$, to the first order in b. Hence to the same order

$$pV = NkT\left(1+\frac{b}{V}\right).$$

This method clearly leaves much to be desired on the score of rigour; moreover it is not easily susceptible of extension to closer approximations. Boltzmann (*loc. cit.*) obtains a second approximation by a similar direct calculation†; but his result is perhaps more safely based on the general method developed below.

§ 1. The phase-space available (considering only the positional coordinates) is that part of a $3N$-dimensional volume Ω (whose boundary is determined by that of the gas itself) for which no one of a set of conditions, in number $\frac{1}{2}N(N-1)$, of the form

$$(x_r-x_s)^2+(y_r-y_s)^2+(z_r-z_s)^2 \geqslant D^2 \ldots\ldots\ldots\ldots(1)$$

* Cf. Boltzmann, *Vorlesungen über Gastheorie*, II, § 61.

† Of the phase-integral, not of i, I separately.

Digitized by Google

is violated. Suppose we choose a set of k of these conditions, and calculate the volume of that part of Ω in which they are all violated; let Ω_k denote the sum of this quantity for all possible sets of conditions. Then

$$I = \Omega - \Omega_1 + \Omega_2 - \ldots\ldots + (-1)^r \Omega_r + \ldots \ldots\ldots (2).$$

For an element $d\tau$ of Ω in which s conditions are violated is counted ${}_sC_r$ times in $\Omega_r (r \leqslant s)$; and

$$\sum_{r=0}^{s} (-1)^r \cdot {}_sC_r = (1-1)^s = 0 \text{ if } s > 0$$
$$= 1 \qquad \text{if } s = 0.$$

Hence the sum on the right-hand side of (2) is precisely the volume of that portion of Ω in which $s = 0$. Jeans* adopts this method and finds

$$\Omega = V^N, \ \Omega_1 = \tfrac{1}{2} N(N-1) \cdot \tfrac{4}{3}\pi D^3 \cdot V^{N-1}.$$

Treating Ω_1 as a first order correction to Ω and calculating i similarly he deduces the value of p found above. He does not notice that Ω_1/Ω increases indefinitely with N, whatever the density, and that it is in fact very large for any observable mass of gas. It is a little surprising, but really might have been expected, that the terms in (2) should increase in absolute value at first; for the number of pairs of molecules engaged in a close encounter at any particular time is very large, so that s also will be very large for nearly the whole of Ω. There is therefore no alternative but to formulate the series (2) exactly, and then devise some means of summing it.

To illustrate the method we first suppose only one kind of molecule present, and that the r conditions violated in any term of Ω_r involve $2r$ different molecules. The number of ways of selecting r pairs of molecules is

$$\frac{N!}{2^r r! \, (N-2r)!},$$

hence
$$\Omega_r = \frac{N!}{2^r \cdot r! \, (N-2r)!} V^{N-r} (\tfrac{4}{3}\pi D^3)^r.$$

Writing $-x$ for $\frac{2}{3} N\pi D^3 / V$ (b in the usual notation) we find

$$I = V^N F(x) = V^N \sum_0 \frac{N!}{r! \, (N-2r)!} \frac{x^r}{N^r}.$$

If we also use y for $x^{-\frac{1}{2}}$, we have

$$N \frac{dF}{dx} = y^{2-N} \frac{d^2}{dy^2} (y^N F)$$
$$= N(N-1) F - 2(2N-3) x \frac{dF}{dx} + 4x^2 \frac{d^2F}{dx^2}.$$

* *Dynamical Theory of Gases*, p. 158.

Digitized by Google

F is the solution of this equation which is regular at $x = 0$ and takes the value 1 there. Put $F = G^N$; then

$$\frac{1}{G}\frac{dG}{dx} = \left(1 - \frac{1}{N}\right) - \left(4 - \frac{6}{N}\right)\frac{x}{G}\frac{dG}{dx} + 4\left[\left(1 - \frac{1}{N}\right)\left(\frac{x}{G}\frac{dG}{dx}\right)^2 + \frac{x^2}{NG}\frac{d^2G}{dx^2}\right].$$

Now N is very large; hence we may write in place of the above

$$\frac{1}{G}\frac{dG}{dx} = 1 - \frac{4x}{G}\frac{dG}{dx} + 4\left(\frac{x}{G}\frac{dG}{dx}\right)^2,$$

or, if $g = \log G$,

$$\frac{dg}{dx} = \left(1 - 2x\frac{dg}{dx}\right)^2 \qquad \text{.....(3).}$$

This equation can be solved without difficulty in finite terms, though of course no physical importance attaches to any but the terms of the first order in x. We find

$$\frac{dg}{dx} = \frac{1 + 4x - \sqrt{(1 + 8x)}}{8x^2}$$

and

$$g = \log\frac{1 + \sqrt{(1 + 8x)}}{2} - \frac{1 + 4x - \sqrt{(1 + 8x)}}{8x}.$$

If this were of general application the value of $-x$ or b could not exceed $\frac{1}{8}$; but this is not the case. We have of course

$$g = x + O(x^2);$$

hence Ng is large and negative, and $I = V^N e^{Ng}$ negligibly small compared with V^N.

§ 2. To evaluate I in general we must first expand each Ω_r. Consider a group of a molecules; their positions can be represented in a space of $3a$ dimensions. At any particular point of this space a certain set of conditions (1) are violated, and there is a whole region in which these same conditions are all broken. We can enumerate all the possible sets of conditions (1) such that each set binds the molecules it involves into a single connected group, the connecting links being the conditions (1). For any such set of conditions and the corresponding connected group of molecules there will be a symmetry number σ which is the number of permutations of the molecules among themselves which leave the set of conditions unaltered. In the $3a$-dimensional space there is a definite region in which these conditions (and possibly others as well) are violated; by permuting the molecules among themselves we get $(a!/\sigma)$ such regions, which may of course overlap. We now enumerate the *types* of such sets of conditions, and introduce the

Digitized by Google

following notation: in the rth type the number of conditions is p_r, and the number of molecules involved is a_r; we write ϵ_r for $(-1)^{p_r}$ and define a quantity ξ_r so that

$$\epsilon_r \xi_r V^{a_r} \sigma_r / N^{a_r - 1}$$

is the volume of the region of a $3a_r$-dimensional space corresponding to the particular set of conditions of the rth type.

In accordance with the definition of Ω_k we now choose any k conditions (1) and calculate the volume of phase-space in which they are violated. Any such set resolves into a number of sets of the types defined above, the corresponding groups of molecules being mutually exclusive; let there be ν_r sets of the rth type, so that

$$k = \Sigma \nu_r p_r \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4).$$

The number of sets of conditions for which the ν_r have assigned values is

$$\frac{N!}{(N - \Sigma a_r \nu_r)!\ \Pi (\nu_r!\ \sigma_r^{\nu_r})}.$$

Hence

$$\Omega_k = \Sigma \frac{N!\ V^{N - \Sigma a_r \nu_r}}{(N - \Sigma a_r \nu_r)!\ \Pi (\nu_r!\ \sigma_r^{\nu_r})} \Pi \left(\frac{\epsilon_r \xi_r V^{a_r} \sigma_r}{N^{a_r - 1}} \right)^{\nu_r},$$

the summation being over all positive integral values of the ν_r satisfying (4). Thus

$$I = \Sigma (-1)^k \Omega_k = V^N \Sigma \frac{N!\ \Pi \xi_r^{\nu_r}}{(N - \Sigma a_r \nu_r)!\ \Pi \nu_r!} N^{\Sigma \nu_r - \Sigma a_r \nu_r} \quad \ldots (5),$$

the summation being over all positive integral ν_r. Clearly ξ_r is $O(b^{a_r - 1})$ in the usual notation.

Write

$$F = F(\xi_r, y) = \Sigma \frac{(yN)^{N - \Sigma a_r \nu_r} \Pi (\xi_r N)^{\nu_r}}{(N - \Sigma a_r \nu_r)!\ \Pi \nu_r!} N!\ N^{-N}.$$

Then

$$\frac{1}{N} \frac{\partial F}{\partial \xi_r} = \left(\frac{1}{N} \frac{\partial}{\partial y} \right)^{a_r} F.$$

Moreover F is defined completely by these equations together with the conditions of being regular in the (ξ_r) at the origin and taking the value y^N there. Now if $a_r = a_s$, that is, if the number of molecules concerned in the rth and sth types of connected groups is the same, we have

$$\frac{\partial F}{\partial \xi_r} = \frac{\partial F}{\partial \xi_s},$$

and therefore F is a function of the sums $\Sigma \xi_{t\,(a_t = r)}$ only. Writing x_r for these we have

$$F = \Sigma \frac{(yN)^{N - \Sigma r \nu_r} \Pi (x_r N)^{\nu_r}}{(N - \Sigma r \nu_r)!\ \Pi \nu_r!} N!\ N^{-N},$$

Digitized by Google

and $$\frac{1}{N}\frac{\partial F}{\partial x_r} = \left(\frac{1}{N}\frac{\partial}{\partial y}\right)^r F.$$

Put $$y = e^z, \quad \delta = \Sigma r x_r \frac{\partial}{\partial x_r};$$

then $$\left(\frac{\partial}{\partial z} + \delta\right)^k F = N^k F.$$

Also write $$F = e^{Ng};$$

then

$$\frac{1}{N}\frac{\partial F}{\partial x_r} = \left(\frac{1}{N}e^{-z}\frac{\partial}{\partial z}\right)^r F$$

$$= e^{-rz}\left(\frac{1}{N}\frac{\partial}{\partial z} - \frac{r-1}{N}\right)\ldots\ldots\left(\frac{1}{N}\frac{\partial}{\partial z} - \frac{1}{N}\right)\left(\frac{1}{N}\frac{\partial}{\partial z}\right)F$$

$$= e^{-rz}\left(1 - \frac{r-1}{N} - \frac{\delta}{N}\right)\ldots\ldots\left(1 - \frac{1}{N} - \frac{\delta}{N}\right)\left(1 - \frac{\delta}{N}\right)F.$$

Now put $y = 1$ or $z = 0$; we find

$$\frac{1}{N}\frac{\partial F}{\partial x_r} = \left(1 - \frac{r-1}{N} - \frac{\delta}{N}\right)\ldots\ldots\left(1 - \frac{1}{N} - \frac{\delta}{N}\right)F(1 - \delta g)$$

$$= \left(1 - \frac{r-1}{N} - \frac{\delta}{N}\right)\ldots\ldots\left(1 - \frac{2}{N} - \frac{\delta}{N}\right)$$

$$\times F\left(1 - \frac{1}{N} - \delta g - \frac{\delta}{N}\right)(1 - \delta g)$$

$$= F\left(1 - \frac{r-1}{N} - \delta g - \frac{\delta}{N}\right)\ldots\ldots\left(1 - \frac{1}{N} - \delta g - \frac{\delta}{N}\right)(1 - \delta g).$$

Making $N \to \infty$ we get

$$\frac{\partial g}{\partial x_r} = (1 - \delta g)^r = h^r \quad \ldots\ldots\ldots\ldots\ldots\ldots (6),$$

where h is written for $(1 - \delta g)$. h satisfies the relation

$$h = 1 - \Sigma r x_r h^r \quad \ldots\ldots\ldots\ldots\ldots\ldots (7).$$

It is easily verified that subject to (7) the partial differential equations (6) are integrable. It is now clear how to evaluate g in series. We first solve (7) for h by successive approximations and then integrate the equations (6), putting $g = 0$ at the origin $(x_r = 0)$. It is then given by

$$I = V^N e^{Ng}.$$

The method extends at once to a gas composed of several kinds of molecules. It is sufficient to consider a mixture of two kinds only, their numbers being N_a, N_b. We now define ξ_r so that the

Digitized by Google

volume of a $3(a_r + b_r)$-dimensional space corresponding to the rth kind of group (composed of a_r molecules of the first kind and b_r of the second kind) is

$$\epsilon_r \xi_r V^{a_r+b_r} \sigma_r / N_0^{a_r+b_r-1}.$$

Here N_0 is supposed to be a number of the order of magnitude of N_a, N_b; there is some advantage in making it either an absolute constant, such as Loschmidt's number, or proportional to V. We now find

$$I = V^{N_a+N_b} F(\xi_r, 1, 1),$$

where

$$F(\xi_r, y, z) = \Sigma \frac{y^{N_a - \Sigma a_r \nu_r} z^{N_b - \Sigma b_r \nu_r} \Pi \xi_r^{\nu_r}}{(N_a - \Sigma a_r \nu_r)!\,(N_b - \Sigma b_r \nu_r)!\,\Pi \nu_r!} N_0^{\Sigma \nu_r (a_r+b_r-1)} N_a!\, N_b!,$$

and hence

$$\frac{1}{N_0}\frac{\partial F}{\partial \xi_r} = \left(\frac{1}{N_0}\frac{\partial}{\partial y}\right)^{a_r}\left(\frac{1}{N_0}\frac{\partial}{\partial y}\right)^{b_r} F.$$

We therefore put $x_{r,s} = \Sigma \xi_{t\,(a_t=r,\, b_t=s)}$

and making the same transformations as before we get finally

$$\frac{\partial g}{\partial x_{r,s}} = \left(\frac{N_a}{N_0} - \delta_a g\right)^r \left(\frac{N_b}{N_0} - \delta_b g\right)^s = h_a{}^r h_b{}^s \quad \ldots\ldots(8),$$

where

$$\left.\begin{aligned} h_a &= \nu_a - \delta_a g = \nu_a - \Sigma r x_{r,s} h_a{}^r h_b{}^s \\ h_b &= \nu_b - \delta_b g = \nu_b - \Sigma s x_{r,s} h_a{}^r h_b{}^s \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots(9).$$

Here ν_a, ν_b, written for N_a/N_0, N_b/N_0 respectively, vary as the densities of the two components, and g vanishes at the origin ($x_{r,s} = 0$). It should be noted that in (7) the summation begins at $r = 2$, and in (9) the terms $r = 1$, $s = 0$ and $r = 0$, $s = 1$ are absent from the summations. These terms are precisely the terms appearing on the left-hand sides of these equations, which take a more symmetrical form on transposition of the series; but the form in which they are written is that adapted to their actual use.

§ 3. We shall now calculate the first few terms in the expansion of g, remembering that x_r is $O(b^{r-1})$. For a pure gas we have

$$h = 1 - \Sigma r x_r h^r, \quad h^r = \partial g / \partial x_r.$$

We obtain successively

$$h = 1,$$
$$h = 1 - 2x_2,$$
$$h = 1 - 2x_2(1 - 4x_2) - 3x_3,$$

giving

$$g = x_2,$$
$$g = x_2 - 2x_2^2 + x_3,$$
$$g = x_2 + (x_3 - 2x_2^2) + (x_4 - 6x_2 x_3 + \tfrac{20}{3} x_2^3).$$

Digitized by Google

For a mixture we have

$$h_a = \nu_a - \Sigma r x_{r,s} h_a^r h_b^s,$$

$$h_b = \nu_b - \Sigma s x_{r,s} h_a^r h_b^s,$$

giving successively

$$h_a = \nu_a, \quad h_b = \nu_b,$$

$$h_a = \nu_a (1 - 2x_{2,0}\nu_a - x_{1,1}\nu_b), \quad h_b = \nu_b (1 - 2x_{0,2}\nu_b - x_{1,1}\nu_a),$$

$$\begin{aligned} h_a = \nu_a &- 2x_{2,0}\nu_a^2 (1 - 4x_{2,0}\nu_a - 2x_{1,1}\nu_b) \\ &- x_{1,1}\nu_a\nu_b [1 - 2x_{2,0}\nu_a - x_{1,1}(\nu_a + \nu_b) - 2x_{0,2}\nu_b] \\ &- 3x_{3,0}\nu_a^3 - 2x_{2,1}\nu_a^2\nu_b - x_{1,2}\nu_a\nu_b^2, \end{aligned}$$

with a corresponding formula for h_b. The terms in g of the first two orders are therefore

$$\begin{aligned} g = &(\nu_a^2 x_{2,0} + \nu_a\nu_b x_{1,1} + \nu_b^2 x_{0,2}) \\ &+ \nu_a^3 (x_{3,0} - 2x_{2,0}^2) + \nu_a^2\nu_b (x_{2,1} - 2x_{2,0}x_{1,1} - \tfrac{1}{2}x_{1,1}^2) \\ &+ \nu_a\nu_b^2 (x_{1,2} - 2x_{0,2}x_{1,1} - \tfrac{1}{2}x_{1,1}^2) + \nu_b^3 (x_{0,3} - 2x_{0,2}^2) \ldots (10). \end{aligned}$$

I find for the coefficient of $\nu_a^3\nu_b$

$$x_{3,1} - (4x_{2,0} + x_{1,1})\, x_{2,1} + \tfrac{1}{3}x_{1,1}^3 + 2x_{2,0}x_{1,1}^2 + 8x_{2,0}^2 x_{1,1},$$

and for that of $\nu_a^2\nu_b^2$

$$\begin{aligned} x_{2,2} - 2x_{2,1}(x_{0,2} + x_{1,1}) - 2x_{1,2}(x_{2,0} + x_{1,1}) \\ + (3x_{2,0} + 3x_{0,2} + x_{1,1})\, x_{1,1}^2 + 4x_{2,0}x_{1,1}x_{0,2}. \end{aligned}$$

That of $\nu_a\nu_b^3$ can be written down from symmetry, and those of ν_a^4, ν_b^4 from the result for a pure gas.

We shall now calculate x_2, x_3, $x_{2,0}$, etc., $x_{3,0}$, etc. for the case of hard spherical molecules. x_2 is a single ξ, and

$$p = 1, \quad \sigma = 2.$$

Hence
$$x_2 = -\frac{N}{2V^2} V \cdot \tfrac{4}{3}\pi D^3 = -\frac{\frac{2}{3}N\pi D^3}{V},$$

D being the diameter of a molecule. For a gaseous mixture we find

$$x_{2,0} = -\frac{\frac{2}{3}N_0\pi D_a^3}{V}, \quad x_{0,2} = -\frac{\frac{2}{3}N_0\pi D_b^3}{V},$$

$$x_{1,1} = -\frac{\frac{4}{3}N_0\pi D_{ab}^3}{V},$$

D_a, D_b being the respective diameters and D_{ab} the sum of the radii.

Again, x_3 is the sum of two different ξ's. For the first one

Digitized by Google

molecule overlaps two others, but these do not necessarily overlap one another. Hence $p = 2$, $\sigma = 2$, and

$$\frac{2\xi V^3}{N^2} = V(\tfrac{4}{3}\pi D^3)^2 \text{ or } \xi = 2x_2^2.$$

For the other ξ each pair of molecules overlap, and we have $p = 3$, $\sigma = 6$. When the first two have their centres at a distance $r < D$, the volume in which the centre of the third must lie is

$$2K(r) = 2\int_{\frac{1}{2}r}^{D} \pi(D^2 - x^2)\,dx = 2\pi\left(\tfrac{2}{3}D^3 - \tfrac{1}{2}D^2 r + \frac{r^3}{24}\right).$$

Hence
$$-\frac{6\xi V^3}{N^2} = V\int_0^D 4\pi r^2 . 2K(r)\,dr,$$

or
$$\xi = -\frac{N^2}{6V^2} . 8\pi^2 \int_0^D r^2\left(\tfrac{2}{3}D^3 - \tfrac{1}{2}D^2 r + \frac{r^3}{24}\right)dr$$

$$= -\frac{4\pi^2 N^2}{3V^2} D^6 [\tfrac{2}{9} - \tfrac{1}{8} + \tfrac{1}{144}]$$

$$= -\frac{5\pi^2 N^2 D^6}{36 V^2} = -\tfrac{5}{16} x_2^2.$$

Thus
$$x_3 = (2 - \tfrac{5}{16})\, x_2^2.$$

For a mixture we have of course

$$x_{3,0} = (2 - \tfrac{5}{16})\, x_{2,0}^2,$$

just as for a pure gas. $x_{2,1}$ is the sum of three terms. For the first one the b-molecule overlaps each of the a-molecules; we have $p = 2$, $\sigma = 2$, and

$$\xi = \tfrac{1}{2} x_{1,1}^2.$$

For the second term one of the a-molecules overlaps each of the others; hence $p = 2$, $\sigma = 1$, and

$$\xi = 2x_{2,0} x_{1,1}.$$

For the third term each pair of molecules overlap, and $p = 3$, $\sigma = 2$. When the a-centres are at distance $r < D_a$, that of the b-molecule must lie in a volume

$$2K(r) = 2\int_{\frac{1}{2}r}^{D_{ab}} \pi(D_{ab}^2 - x^2)\,dx.$$

Hence

$$\xi = -\frac{N_0^2}{2V^2}\int_0^{D_a} 4\pi r^2 . 2K(r)\,dr$$

$$= -\frac{4\pi^2 N_0^2}{V^2}\int_0^{D_a} r^2\left(\tfrac{2}{3}D_{ab}^3 - \tfrac{1}{2}rD_{ab}^2 + \frac{r^3}{24}\right)dr$$

$$= -4\pi^2 N_0^2 V^{-2} D_a^3 [\tfrac{2}{9} D_{ab}^3 - \tfrac{1}{8} D_a D_{ab}^2 + \tfrac{1}{144} D_a^3].$$

Digitized by Google

Inserting in the expansion for g we find that the second order terms expressed in terms of $x_{2,0}$, $x_{1,1}$, $x_{0,2}$ cancel out and we are left with

$$g = (\nu_a^2 x_{2,0} + \nu_a \nu_b x_{1,1} + \nu_b^2 x_{0,2}) - \tfrac{5}{16}(\nu_a^3 x_{2,0}^2 + \nu_b^3 x_{0,2}^2)$$
$$- 4\pi^2 N_0^2 D_a^3 V^{-2} \nu_a^2 \nu_b [\tfrac{2}{9} D_{ab}^3 - \tfrac{1}{8} D_{ab}^2 D_a + \tfrac{1}{144} D_a^3]$$
$$- 4\pi^2 N_0^2 D_b^3 V^{-2} \nu_a \nu_b^2 [\tfrac{2}{9} D_{ab}^3 - \tfrac{1}{8} D_{ab}^2 D_b + \tfrac{1}{144} D_b^3] \ldots (10').$$

§ 4. In the above, difficulties connected with the boundary have not been mentioned. As a matter of fact they take care of themselves if the numbers ξ_r, x_r, $x_{r,s}$ are evaluated strictly as defined. We may if we wish take them explicitly into account by writing $x_r + x_r'$ instead of x_r, where x_r is evaluated as above, neglecting the boundary, so that

$$x_r (V/N)^{r-1}$$

is a function only of r and the diameter D of a molecule, while x_r' depends on the size and shape of the boundary. It would be effectively proportional to the area (S, say) of the latter, and we should have

$$x_r'/x_r = O(SD/V) = O(x_2/N)^{\frac{1}{3}} = O(b/N)^{\frac{1}{3}}.$$

Hence the effect of such considerations on the value of g is negligible. It is easily seen that the method of considering them described is equivalent to regarding the boundary as another kind of "molecule," of which only one is present. In the same way we can deal with the case in which a particular molecule is assigned a definite position; if the latter is in the interior of the gas we evidently have

$$x_r' = (r+1)\, x_{r+1}/N.$$

In such a way we can calculate i and hence the pressure.

Unfortunately however the calculation is open to the same objections as the original one. For to find a ratio such as i/I accurately, say to order b^r, we need the value of g in each case correct to b^r/N; and this is precisely what we do not know.

It would presumably be possible to evade this difficulty by including in the differential equations the terms in N^{-1}, which would be linear in the differential coefficients of the second order; we could then obtain the terms in N^{-1} in g.

There is however another method of deducing the equation of state of an assembly from statistical theory, a method of extraordinary power. Due originally to Gibbs, its fundamental character has been emphasised in recent times by applications and extensions of it to quantised systems*. It will be remembered that the

* Darwin and Fowler, *Phil. Mag.*, vol. XLIV, and other papers.

Digitized by Google

definition of I was somewhat complex; in the method referred to there is introduced the "phase-integral"

$$H(\mathfrak{H}) = \int I(E)\,\mathfrak{H}^E dE = \int \mathfrak{H}^E d\tau \quad \ldots\ldots\ldots\ldots(11).$$

It is found that all the equilibrium properties of the assembly can be obtained from $H(\mathfrak{H})$. In fact in the quantum theory case $H(\mathfrak{H})$ is a power series (of a somewhat generalised kind) and it is natural to suppose that if we know it we can deduce $I(E)$. The phase-integral is dominated completely by the maximum of the integrand (in the first form given above); this occurs when

$$\frac{1}{I}\frac{dI}{dE} + \log \mathfrak{H} = 0 \quad \ldots\ldots\ldots\ldots(12).$$

This equation determines a value of the parameter $\mathfrak{H}$ which is characteristic of the state of the assembly when it has energy E; it may be shown that

$$\mathfrak{H} = e^{-1/kT} \ldots\ldots\ldots\ldots(13),$$

where T is the absolute temperature. If Y be a function satisfying certain conditions, fulfilled by all measurable quantities, its average value is given by

$$\overline{Y} = \frac{H_Y(\mathfrak{H})}{H(\mathfrak{H})} = \int Y \mathfrak{H}^E d\tau \Big/ \int \mathfrak{H}^E d\tau \quad \ldots\ldots\ldots\ldots(14),$$

with a proportional error of the order N^{-1}. In particular

$$E = \frac{H_E(\mathfrak{H})}{H(\mathfrak{H})} = \mathfrak{H}\frac{\partial}{\partial \mathfrak{H}} \log H(\mathfrak{H}) \ldots\ldots\ldots\ldots(15),$$

while if Y be the reaction $(-\partial E/\partial y)$ corresponding to any dynamically measurable coordinate y we have

$$\overline{Y} = -\frac{1}{\log \mathfrak{H}}\frac{\partial}{\partial y} \log H(\mathfrak{H}) \quad \ldots\ldots\ldots\ldots(16).$$

Hence the pressure is given by

$$p = kT\frac{\partial}{\partial V} \log H(\mathfrak{H}) \quad \ldots\ldots\ldots\ldots(17).$$

§ 5. Now our method of calculating I for a gas of impenetrable molecules extends at once to the evaluation of $H(\mathfrak{H})$ for any gas. The impulse coordinates again give a factor in $H(\mathfrak{H})$ entirely independent of the rest; for the remaining factor the integrand is $\mathfrak{H}^E$, where E now represents the total potential energy of the assembly. This will be roughly a sum of terms corresponding to the various groups of molecules engaged at any time in a close encounter; we obtain successive approximations to the appropriate value of E and hence of $\mathfrak{H}^E$ at any point in the generalised space

Digitized by Google

by making the groups of molecules we regard as independent more and more all-inclusive. We find

$$H(\vartheta) = V^N e^{Ng} \quad \text{.....(18),}$$

where in g the value of x_r is now such that $x_r V^r N^{1-r} r!$ is the integral taken over a $3r$-dimensional space of a quantity we may denote by $u_{(r)}$. For $r=2$ it is

$$u_{(2)} = \vartheta^{E_{12}} - 1 = \vartheta_{(2)} - 1, \text{ say;}$$

for $r=3$, $\quad u_{(3)} = \vartheta_{(3)} - \Sigma u_{(2)} - 1 = \vartheta_{(3)} - \Sigma\vartheta_{(2)} + 2.$

In general we must have

$$\vartheta_{(n)} = \Sigma\Pi u_{(i_r)} \quad \text{.....(19),}$$

where the summation is taken over all possible divisions of the n molecules into distinct groups containing say $i_1, i_2, \ldots i_s$ molecules respectively. If we fix our attention on a particular molecule, say the nth, we can sum all the terms in which the group containing this molecule is the same; this leads to

$$\vartheta_{(n)} = u_{(n)} + \Sigma u_{(n-1)}\vartheta_{(1)} + \Sigma u_{(n-2)}\vartheta_{(2)} + \ldots \quad \text{.....(20),}$$

where the u's refer to the various possible sub-groups which include the nth molecule, the ϑ's to the complementary sub-groups. $\vartheta_{(1)} = 1$ always. Now if we denote by

$$(i_1, i_2, \ldots\ldots i_s)$$

the coefficient of $\vartheta_{(i_1)}\vartheta_{(i_2)} \ldots\ldots \vartheta_{(i_s)}$ in $u_{(\Sigma i_r)}$, we get at once

$$0 = (i_1, i_2, \ldots i_s) + (i_2, \ldots i_s) + (i_1, i_3, \ldots i_s) + \ldots + (i_1, \ldots i_{s-2}, i_s),$$

supposing the sth group to be the one containing the nth molecule. The above equation fails only when $s=1$, when it is replaced by $1 = (n)$. Hence

$$(i_1, i_2, \ldots\ldots i_s) = (-1)^{s-1}(s-1)!$$

and so $\quad u_{(n)} = \Sigma(-1)^{s-1}(s-1)!\,\vartheta_{(i_1)}\vartheta_{(i_2)}\vartheta_{(i_3)} \ldots\ldots \vartheta_{(i_s)} \ldots (21),$

the summation being taken over every possible division of the n molecules into distinct sub-groups. For example,

$$u_{(4)} = \vartheta_{1234} - (\Sigma\vartheta_{123} + \Sigma\vartheta_{12}\vartheta_{34}) + 2\Sigma\vartheta_{12} - 6.$$

ϑ_{1234} now denotes $\vartheta^{E_{1234}}$, and E_{1234} the mutual potential energy of the four molecules, all others being supposed absent. We have omitted the factors $\vartheta_1, \vartheta_2, \vartheta_3, \vartheta_4$, which are all unity in the case considered so far.

When the gas is in a field of force, (21) still holds good, but we no longer have $\vartheta_{(1)} = 1$; moreover the first term of the series will not be V^N but V'^N, where

$$V' = \iiint \vartheta_{(1)}\, dx\,dy\,dz = \iiint \vartheta^E\, dx\,dy\,dz,$$

Digitized by Google

E being the potential energy of a single molecule in the field of force. It is therefore necessary to re-define x_r, using V' in place of V.

For a gaseous mixture we have a V_a and a V_b which are in general different; we find

$$H(\mathfrak{Z}) = V_a^{N_a} V_b^{N_b} e^{N_0 g},$$

where, in g, $x_{r,s}$ is such that

$$x_{r,s} V_a^r V_b^s N_0^{1-s-r} r!\ s! = \int u_{(r+s)}\, d\tau_{r,s}.$$

The formal connection of the u's with the $\mathfrak{Z}$'s, i.e. equation (21), is of course unaffected by distinctions between different kinds of molecules; but the $\mathfrak{Z}$'s themselves will be affected by such distinctions. We can no longer say positively what the orders of x_r, $x_{r,s}$ must be; what we can say however is that in the simple case x_r varies as the $(r-1)$th power of the density, and that in the general case $x_{r,s}$ occurs in g with a coefficient $\nu_a^r \nu_b^s$.

To calculate the pressure we note that

$$V_a \propto V, \quad V_b \propto V,$$

and
$$\log H(\mathfrak{Z}) = \text{const.} + (N_a + N_b) \log V + N_0 g.$$

Also
$$x_{r,s} \propto V^{r+s-1}.$$

Hence
$$p = kT\left(\frac{N_a + N_b}{V} + N_0 \frac{\partial g}{\partial V}\right)$$
$$= N_0 kT\left[(\nu_a + \nu_b) + g'\right] V^{-1} \quad \ldots\ldots\ldots\ldots\ldots\ldots (22),$$

where g' is obtained from g by multiplying by $(r+s-1)$ the terms in $\nu_a^r \nu_b^s$. This formula applies of course in all cases; we have no need of the doubtful method of i/I.

In conclusion I would like to say how much this paper owes to the influence of Mr R. H. Fowler.

Digitized by Google

The limits of classical scattering. By P. M. S. BLACKETT, M.A., Fellow of King's College, Cambridge.

[*Received* 5 February, *read* 14 March 1927.]

The methods of geometrical optics provide an approximate solution of the equation of wave propagation, if the following condition, given by De Broglie*, is satisfied. If V is the phase velocity, λ the wave-length, l the direction of greatest increase of V, and θ the angle between l and V, then the condition is that

$$Z \equiv \frac{\cos\theta}{2\pi} \frac{1}{V} \frac{dV}{dl} \lambda << 1 \quad \text{.....(1),}$$

that is, the relative change of V over a distance of the order of λ must be small†.

If this condition is satisfied for the waves associated by De Broglie's theory with the motion of a material particle, the particle will move according to the laws of classical mechanics. For example, a charged particle will move classically in a hyperbolic orbit round another fixed charge only if $Z << 1$ at all points of the orbit. Wentzel‡ has recently shown that classical inverse square law scattering can in fact be obtained as an approximate solution of Schrödinger's differential equation. If this condition is not satisfied a rigorous solution of the wave equations is required to describe the phenomena. It is to be expected that this solution will correspond to some kind of diffraction effect.

Let a particle of mass m and charge e_1 and initial velocity v_0 move in a hyperbolic orbit in the repulsive field of a fixed charge e_2. If 2α is the angle between the asymptotes, the equation of the orbit in the tangential polar coordinates r, p is

$$\frac{l}{p^2} + \frac{2}{r} = \frac{1}{a} \quad \text{.....(2),}$$

where

$$l = a \tan^2\alpha \quad \text{.....(3),}$$

and

$$a = e_1 e_2 / m v_0^2 \quad \text{.....(4).}$$

The orbit is determined in terms of the striking radius p_0 by the equation

$$p_0 = a \tan\alpha.$$

* De Broglie, *Journ. de Phys.* (1926), p. 322.

† In De Broglie's statement of the condition the direction l is left undefined. For a plane wave it can be proved that l must be taken as the direction of greatest increase of V. I am indebted to Mr L. H. Thomas for pointing out that if the wave is not plane, the condition (1) will contain other terms involving the curvature of the *wavefront*. It seems unlikely that these terms will be of importance in the cases discussed in this paper.

‡ Wentzel, *Zeit. für Phys.*, vol. XL (1926), p. 590.

Digitized by Google

The waves associated with the particle have a length

$$\lambda = h/mv, \quad (\lambda_0 = h/mv_0) \quad \ldots\ldots\ldots\ldots\ldots\ldots(5),$$

and are propagated with a velocity

$$V = c/v \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(6).$$

Since $$v^2 = v_0^2 - 2e_1e_2/mr = v_0^2(1 - 2a/r) \quad \ldots\ldots\ldots\ldots(7),$$

the radius vector r is also the direction of greatest increase of V, so that

$$dV/dl = dV/dr,$$

and $$\sin\theta = p/r \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(8).$$

Thus from (1) and (6) we get

$$Z = \frac{\lambda}{4\pi}\frac{\cos\theta}{v^2}\frac{d(v^2)}{dr},$$

which with (5) gives

$$Z = \frac{\lambda_0}{4\pi}\frac{v_0}{v^3}\cos\theta\frac{d(v^2)}{dr}.$$

Using (2), (4), (7) and (8), we get

$$Z = \lambda_0 a\,(r^2 - 2ar - la)^{\frac{1}{2}}/2\pi r\,(r - 2a)^2.$$

If we write $x = r/a$ and $\rho = l/a$, we get

$$Z = \lambda_0/kl,$$

where $$k = 2\pi x\,(x-2)^2/\rho\,(x^2 - 2x - \rho)^{\frac{1}{2}} \ldots\ldots\ldots\ldots\ldots(9).$$

The condition (1) will be satisfied at all points of the orbit if

$$\frac{\lambda_0}{k_m l} << 1 \ldots\ldots\ldots\ldots\ldots\ldots\ldots(10),$$

where k_m is the minimum value of k for the given orbit. We can find k_m numerically from (9). The values are given below for different values of the deflection $\phi = \pi - 2\alpha$,

$\phi°$	0	20	40	60	80	100	120	140	160	180
k_m	16·1	14·0	12·0	10·4	8·6	6·9	5·0	3·4	1·8	0.

The condition for the classical description of an orbit is then that the wave-length λ_0 corresponding to the initial velocity of the particle should be small compared with k_m times the semi-latus rectum l of the orbit.

In view of (3), (4) and (5) the condition can be written

$$\frac{\lambda_0}{a}\frac{\cot^2\alpha}{k_m} = \frac{hv_0}{e_1e_2}\frac{\cot^2\alpha}{k_m} << 1 \quad \ldots\ldots\ldots\ldots(11).$$

Digitized by Google

It is interesting to compare this condition with the qualitative one mentioned by Schrödinger* and given also by De Broglie as the quantitative equivalent of condition (1), namely, that the wave-length should be small compared with the radius of curvature of the orbit. Now for a hyperbolic orbit minimum radius of curvature occurs at the apse and is numerically equal to the semi-latus rectum l. The condition is then that

$$\lambda/l << 1,$$

or, giving λ its value at the apse,

$$\lambda_0/k'l << 1,$$

where $$k' = \{(\sec \alpha - 1)/(\sec \alpha + 1)\}^{\frac{1}{2}}.$$

This constant k' is found to be very nearly $k_m/16{\cdot}1$, for all values of ϕ. In this particular case Schrödinger's rough condition differs from De Broglie's only by a constant numerical factor†.

It can be shown by using a theorem due to Darwin that condition (11) is also valid in the general case of two particles of comparable mass, if the angle α is replaced by Θ, the angle of projection of the particle initially at rest.

The condition for classical scattering is always satisfied, as it must be, if h is put equal to zero. It is also satisfied when e_1e_2 is large and when v is small (except when α is very nearly zero), that is for the scattering of macroscopic particles. But diffraction effects are to be expected for angles of deflection ϕ sufficiently near 180° even if hv_0/e_1e_2 is very small.

As an example of the application of (11) to the scattering of alpha particles the following cases will be discussed. For an alpha particle of 5 cms. range scattered through 30° by an aluminium nucleus we find $Z \doteqdot 1/100$. Approximate normal scattering is to be expected and is in fact found‡. If the angle of scattering is 90° we find $Z \doteqdot 1/4$ and large deviations from normal scattering are both expected and are found. If the observed deviations in the latter case are in fact due to a diffraction effect we should expect similar deviations in other cases when Z is approximately 1/4. For the collision of an 8·6 cm. alpha particle with a proton, deviations are to be expected at angles of projection about 75°, that is for deflections of the alpha particle of 6°. Such deviations

* Schrödinger, *Ann. d. Phys.* vol. LXXIX (1926), p. 496.

† It should be noticed that though the curvature has its greatest value at the apse, the quantity Z has its maximum some distance away from the apse. Since Z contains the factor $\cos\theta$ it is zero at the apse. There can therefore be no general relation between Schrödinger's and De Broglie's conditions. For instance, for a particle in a circular orbit, $Z = 0$, whatever the size of the orbit.

‡ Bieler, *Proc. Roy. Soc.* A, vol. CV (1924), p. 434.

Digitized by Google

certainly occur*. For the collision of 5 cm. alpha particles with gold nuclei, deviations are to be expected at angles of about 125°. The recent experimental results of Rutherford and Chadwick† are in complete disagreement with this result. Of the many possible causes of this discrepancy two may be mentioned. The deviations observed with aluminium for $Z = 1/4$ may be a *structure* effect; the diffraction effect may not become noticeable till Z is, say, roughly unity, or even greater. The diffraction effect may be present but may not be of a type to be revealed by the measurements; for instance the fairly wide limits of scattering angle in the experiments of Rutherford and Chadwick may have left unrevealed any rapid variations of scattering with angle.

It seems therefore not impossible that the observed deviations from normal scattering, which have hitherto been explained with the aid of classical mechanics by postulating arbitrary deviations from the inverse square law, may now receive an explanation with the aid of the wave mechanics and the inverse square law alone. If other forces connected with the structure of the nucleus, such as polarisation or magnetic forces, have to be taken into account, their effect may complicate, but cannot in general be expected to mask, the diffraction scattering due to the inverse square law alone.

Since the product e_1e_2 but not the mass m enters into (11), the limits of classical scattering of electrons by electrons will be given by this same condition. When the field is attractive, as in the scattering of electrons by nuclei, the condition for classical scattering will still be of the form (11) but k_m will have a different value.

The dimensional part $\lambda/a \propto hv/e_1e_2$ of (11) can be obtained from considerations of dimensions alone. This method also applies when the law of force is of the general type μr^{-n}. From the analogy with optics the required condition is that the ratio λ/L should be small, where L is a length playing the part of the *size* of the diffracting centre. The only dimensional quantities on which L can depend are μ, m and v. It must therefore be of the form

$$(\mu/mv^2)^{1/n-1}.$$

We have therefore $$\lambda/L \propto hv^{\frac{3-n}{n-1}} m^{\frac{2-n}{n-1}} \mu^{\frac{1}{1-n}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (12).$$

For $n = 2$, $\mu = e_1e_2$ and λ/L is proportional to hv/e_1e_2, in agreement with (11). When the forces vary inversely as the cube of the distance, $n = 3$ and λ/L is proportional to

$$h/(m\mu)^{\frac{1}{2}}.$$

Here the velocity does not enter into the expression for λ/L and so the limits of normal scattering should be independent of the velocity.

* Chadwick and Bieler, *Phil. Mag.*, vol. XLII (1921), p. 923.

† Rutherford and Chadwick, *Phil. Mag.*, vol. L (1925), p. 889.

Digitized by Google

When the forces vary inversely as a high power of the distance we approximate to the case of the scattering by a rigid body. Then we have $L = \text{const.}$ and $\lambda/L \propto h/mv$. The deviations from normal scattering now occur when v is small, instead of when v is large, as for the inverse square law. This is due to the fact that in the latter case the effective *size* of the scattering centre of force is proportional to a, and therefore to $1/v^2$, while in the former case the effective *size* is constant.

Elsasser* has suggested that the long free path of very slow electrons observed by Ramsauer can be associated with the decrease of scattering by a rigid obstacle as the wave-length of the scattered wave is increased. Why this effect should only appear with atoms of high symmetry appears clearly from the dimensional argument given above. Elsasser's explanation assumes a diffracting disc whose size is independent of the velocity. This is only obtained when n in (12) is large, that is when the law of force outside the atom varies very rapidly with the distance. Large values of n can only be expected with the most symmetrical atoms.

In optics not only is the condition for geometrical optics that the ratio of the wave-length to a linear dimension of the object should be small, but the intensity in any direction when diffraction occurs is a function of the same ratio. In the dynamical case the numbers of particles scattered in a given direction should therefore be a function of λ/L, that is of the R.H.S. of (12). For the law of the inverse square the *type* of scattering observed should be the same for all cases when v/e_1e_2 has the same value. For the inverse cube law the *type* of diffraction pattern should be independent of the velocity. This may be the origin of the fixed peak in the distribution of scattered electrons from helium, found by Dymond†.

Note added in Proof. In a recent paper (*Göttinger Nachrichten*, 1927) Born has derived the classical scattering formula as an approximate solution of Schrödinger's wave equations, for the scattering of an alpha particle by a neutral hydrogen atom. In both Wentzel's and Born's results deviations from the inverse square law scattering occur for low velocity particles. This appears to be connected with the screening effect of the electron in the hydrogen atom. That normal scattering by a naked nucleus must occur for slow particles, seems certain in view of the occurrence of the product hv in (11) above and in the quantity z_0 in equation (32) of Born's paper. Since we must obtain classical scattering when $h = 0$, we must also obtain classical scattering as v approaches zero.

* Elsasser, *Naturwissenschaften*, vol. XIII (1925), p. 711.
† Dymond, *Nature*, vol. CXVIII (1926), p. 336.

Digitized by Google

On Herschel's Condition and the Optical Cosine Law. By Mr G. C. STEWARD, Gonville and Caius College.

[*Received* 11 February, *read* 28 February 1927.]

§ 1. Sir John Herschel gave the condition which must be satisfied in order that a symmetrical optical system, free from spherical aberration for two conjugate axial points, may also be free from spherical aberration for two neighbouring and conjugate points upon the axis of the system; but Herschel's condition applies only to first order aberration, i.e. to aberration depending upon the cube of the inclination of the ray to the axis. Abbe shewed, later, that this condition could be included in a wider result, viz. that the spherical aberration, supposed zero, is stationary for axial variations provided that the incident and emergent rays for two conjugate axial points, associated with modified magnification m, satisfy the relation

$$m \sin (\theta'/2) = \sin (\theta/2)^*, \qquad \ldots\ldots(1)$$

where θ and θ' are their initial and final inclinations to the axis; and by 'modified' magnification is meant the ratio of the reduced sizes of the image and object.

The form of the condition expressed by (1) brings to mind the well-known sine-condition; which, expressed in the preceding notation, is

$$m \sin \theta' = \sin \theta^*; \qquad \ldots\ldots(2)$$

and the satisfaction of this condition for two conjugate axial points, free from spherical aberration, ensures a good image of a small two-dimensional object placed normal to the axis of the system, at either of the points. In other words, (2) implies the absence of circular coma of *all* orders for this pair of points; and not merely of first order coma†: this has been investigated in preceding papers, where also is considered the precise meaning of a given departure from condition (2). It is desired to investigate here, in a similar way, the meaning of Herschel's condition, expressed by (1); and also the meaning of a given departure from it: and the method adopted depends upon the use of some type of 'characteristic' function associated with the name of Hamilton.

* See also a very elegant proof of these two conditions given by C. Hockin, *J. R. Mic. Soc.* (2), IV (1884).

† "The Aberrations of a Symmetrical Optical System," *Trans. Camb. Phil. Soc.* IX, 1926; "Aberration Diffraction Effects," *Phil. Trans. Roy. Soc.* A, 225, Part II, § 21.

Digitized by Google

It will be observed that in the application both of Herschel's condition and of the sine-condition we have assumed the existence of two conjugate axial points which are free from spherical aberration; but it will appear, in the sequel, that these two conditions, and other conditions also, are themselves included in a much wider result, the 'Optical Cosine Law'*; which contemplates the presence of spherical aberration or indeed of any other type of aberration.

§ 2. Consider a symmetrical optical system, and take parallel rectangular axes of co-ordinates $O(x, y, z)$ and $E'(x', y', z')$ in the object space and in the image space respectively, the axes of x and of x' being coincident with the axis of symmetry; and let O and E' be axial points associated respectively with reduced magnifications m and s. Let a ray, incident and emergent, intersect the planes y, z and y', z' in $P(y, z)$ and $Q(y', z')$; then the reduced path, $V(PQ)$, from P to Q, measured along the ray, will be a function of the four variables y, z, y' and z': from the symmetry of the system these may be replaced by the three variables θ, ϕ and ψ, given by the relations

$$\theta d^2 = y^2 + z^2,\ \phi d^2 = 2(yy' + zz') \text{ and } \psi d^2 = y'^2 + z'^2, \quad (1)$$

where d is the distance from E' to the point, upon the axis, conjugate to O, i.e. $d = s - m$. Here d is the 'modified' and reduced distance, i.e. the reduced distance multiplied by the modified power $J(\equiv K/\mu\mu')$ of the system, μ and μ' being the indices of the end media and K the customary power of the system.

For a system giving perfect definition over the conjugate normal planes associated with magnification m, the reduced path from the point P to its conjugate P', measured along a ray passing through both points, would depend only upon the position of P (or of P'), i.e. would depend only upon the variable θ; so that V, the reduced path from P to Q', would be given by

$$V = f(\theta) - d\sqrt{1 + u}, \qquad \ldots\ldots(2)$$

where $u = \theta - \phi + \psi$ and $f(\theta)$ is an arbitrary function of the variable θ. Any deviation from this result will imply the presence of aberrations for the conjugate planes under consideration; so that if, in any particular case, we have

$$V = f(\theta) - d\sqrt{1 + u} + F, \qquad \ldots\ldots(3)$$

F will be a function of the variables θ, ϕ and ψ, and will give *all* the aberrations of the system, and will depend upon these aberrations alone: F may therefore be named the 'aberration function.'

* T. Smith, "The Optical Cosine Law," *Trans. Opt. Soc.* Vol. xxiv (1922–23), No. 1.

Digitized by Google

The form of the function $f(\theta)$, so far arbitrary, may be chosen to satisfy certain other conditions; which usually will be the absence of spherical aberration of all orders for the pair of conjugate axial points associated with magnification s; and these then may be taken as the axial points of the pupil planes. A system is only to be regarded as fully corrected when aberrations are absent for the magnification m and also this latter condition is satisfied for the entrance and exit pupils. Moreover, very frequently, in cases of practical importance, we may write $s = 1$.

F is a function of θ, ϕ and ψ, the lowest terms in these variables being of the second degree, since the aberrations of the system are essentially of higher order than the first in θ, ϕ and ψ; also for an axial object point $\theta = \phi = 0$, so that the various orders of spherical aberration (in the restricted sense of *axial* aberration)* will depend only upon the terms in ψ, and corresponding to these aberrations we shall have terms in F of the form

$$A_{0,0,n}\psi^n, \qquad \ldots\ldots(4)$$

where n is a positive integer greater than unity.

§ 3. Let now L, M, N and L', M', N' respectively be the direction cosines of an incident ray and the corresponding emergent ray; then from the essential property of the characteristic function V we have

$$L = -\frac{\partial V}{\partial \xi},$$

where ξ is a (reduced) variation, in the object space, parallel to the x-axis of co-ordinates, i.e. parallel to the axis of symmetry of the optical system. But, if ξ be the modified and reduced distance of O from the first principal focus,

$$\xi = 1/m,$$

so that

$$L = m^2 \frac{\partial V}{\partial m};$$

similarly, replacing $1/m$ by $-s$, and so considering the image space, we have

$$L' = -\frac{\partial V}{\partial s}.$$

If now χ be any one of the variables θ, ϕ and ψ we have

$$\frac{\partial \chi}{\partial m} d - 2\chi = 0 \text{ and } \frac{\partial \chi}{\partial s} d + 2\chi = 0;$$

* Cf. *Trans. Camb. Phil. Soc.* IX, 1926, § 29 *et seq.*

Digitized by Google

so that, writing

$$F = \sum_{n=2}^{\infty} A_n \psi^n$$

in (2) § 2, and using the equations for L and L', we have

$$\left.\begin{aligned} L &= m^2 \left(\frac{\partial f}{\partial m}\right)_{\theta=0} + \frac{m^2}{\sqrt{1+\psi}} + m^2 \sum_{n=2}^{\infty} \psi^n \left(nA_n + \frac{\partial A_n}{\partial m}\right) \\ \text{and} \quad L' &= -\frac{1}{\sqrt{1+\psi}} + \sum_{n=2}^{\infty} n\psi^n A_n \end{aligned}\right\}, \quad (1)$$

since A_n is independent of s. In obtaining these results we have assumed an axial object point, i.e. $\theta = \phi = 0$; and $f(0)$ is independent of s. Indeed, if F_1 and F_2 be the principal foci of the system and O' the point conjugate to O, we have

$$F_1O = 1/m \text{ and } F_2O' = -m,$$

so that

$$\begin{aligned} f(0) &= OO' \\ &= F_1F_2 - F_1O + F_2O', \end{aligned}$$

i.e.

$$f(0) = F_1F_2 - 1/m - m;$$

and

$$\frac{\partial f(0)}{\partial m} = 1/m^2 - 1.$$

Thus from (1) above

$$\frac{L-1}{L'-1} = m^2 \frac{1 - (1/\sqrt{1+\psi}) + \Sigma + \Sigma'}{1 - (1/\sqrt{1+\psi}) + \Sigma'}, \quad \ldots\ldots(2)$$

where $\Sigma \equiv \sum_{n=2}^{\infty} n\psi^n A_n$ and $\Sigma' \equiv \sum_{n=2}^{\infty} \psi^n \frac{\partial A_n}{\partial m}$. If now the conjugate axial points O and O' be free from spherical aberration, $A_n = 0$ for all values of n and therefore $\Sigma = 0$; so that from (2), θ and θ' being the initial and final inclinations of the ray to the axis,

$$\frac{\sin(\theta/2)}{\sin(\theta'/2)} = m \left\{1 + \frac{\Sigma'}{1 - (1+\psi)^{-\frac{1}{2}}}\right\}^{\frac{1}{2}}. \quad \ldots\ldots(3)$$

It is clear therefore that if $\frac{\partial A_n}{\partial m} = 0$ for all values of n, i.e. if $\Sigma' = 0$, then the condition

$$\sin(\theta/2) = m \sin(\theta'/2)$$

is satisfied for all rays; and conversely by the expansion of (3) it follows that if this condition be satisfied for all rays, then $\frac{\partial A_n}{\partial m} = 0$ for all values of n. Thus Herschel's condition, applied to two conjugate axial points free from spherical aberration, implies that neighbouring axial points are free from spherical aberration of *all* orders; and conversely.

Digitized by Google

In order to estimate the meaning of a given departure from Herschel's condition, let us write

$$\left(\frac{1}{m}\frac{\sin(\theta/2)}{\sin(\theta'/2)}\right)^2 = 1+\epsilon \qquad \ldots\ldots(4)$$

for any particular ray; then by comparison with (3) we have

$$\epsilon = \frac{\sum_{n=2}^{\infty} A_n'\psi^n}{1-(1/\sqrt{1+\psi})} = (1+\psi+\sqrt{1+\psi})\left(\sum_{n=2}^{\infty} A_n'\psi^n\right),$$

where the dash denotes differentiation with respect to m; so that

$$\epsilon = 2A_2'\psi^2 + (\tfrac{3}{2}A_2' + 2A_3')\psi^3 + \ldots. \qquad \ldots\ldots(5)$$

Considering, therefore, only a first approximation, if the magnification as calculated from Herschel's condition applied to a marginal ray be $m(1+\eta)$, we have

$$\frac{\partial A_2}{\partial m} = \eta\left(\frac{\rho_1}{d}\right)^2,$$

ρ_1 being the radius of the exit pupil; and this gives a measure of the spherical aberration introduced by a slight axial movement of the object plane.

§ 4. The two conditions discussed in the preceding paragraphs are concerned with the attempt to secure point-to-point imagery; it is assumed that we have a pencil of rays, diverging from a point, which after refraction through the optical system converge to a point in the image space: and the condition is sought which will ensure this point-to-point correspondence in the neighbourhood of the base points considered. Two cases have been taken, the base points being upon the axis of the system; we have found the condition that neighbouring axial points should be represented accurately, i.e. Herschel's condition; and we have indicated the condition, the sine-condition, that near points in a plane normal to the axis should be represented perfectly. It is naturally suggested that we attempt to find the condition for perfect representation of *any* given neighbouring point; but this appears more readily, perhaps, from a general theorem to be considered presently. Meanwhile it may be observed that the two conditions already found are, in general, incompatible, i.e. unless $m = \pm 1$; so that it is impossible to secure the perfect representation of a three-dimensional space element. Moreover, three neighbouring points cannot be represented perfectly unless either they lie in a straight line or else upon a plane normal to the axis of the optical system. In the preceding we have omitted, as without practical importance, the one case in which perfect point-to-point correspondence is

Digitized by Google

possible for all points, viz. that in which the optical system is a plane mirror.

A wave surface implies a caustic surface—the corresponding surface of centres; and hitherto the caustic surfaces considered have degenerated into points, so that the corresponding wave surfaces were concentric spheres. And if these wave surfaces are distorted by the presence of 'aberrations,' the conditions discussed can no longer be employed; the caustic surface ceases to be a point. Quite recently a very general theorem has been given by T. Smith under the name of the 'Optical Cosine Law'*; and this law includes as special cases both Herschel's condition and also the sine-condition, as well as other optical laws. The proof given is an analytical one involving the Eikonal of Bruns; the following proof, depending directly upon the 'principle of stationary path,' is given as an alternative. We proceed to prove a necessary lemma.

Lemma. Let $P(x, y, z)$ be a point upon a surface $F(x, y, z) = 0$ and let l, m, n be the direction cosines of some fixed direction; if L, M, N be the direction cosines of the normal at P to the surface, we have

$$\lambda L = \frac{\partial F}{\partial x}, \quad \lambda M = \frac{\partial F}{\partial y}, \quad \lambda N = \frac{\partial F}{\partial z} \quad \text{where} \quad \lambda^2 = \left(\frac{\partial F}{\partial x}\right)^2 + \left(\frac{\partial F}{\partial y}\right)^2 + \left(\frac{\partial F}{\partial z}\right)^2.$$

Let $Q(\xi, \eta, \zeta)$ be a point upon the normal at P to the surface and such that

$$PQ = \rho \cos\theta,$$

ρ being a small constant quantity and θ the inclination of the normal to the direction l, m, n; so that

$$\cos\theta = Ll + Mm + Nn.$$

The equation of the locus of Q is therefore

$$F(\xi - L\rho\cos\theta, \eta - M\rho\cos\theta, \zeta - N\rho\cos\theta) = 0,$$

i.e. to the first approximation it is

$$F(\xi, \eta, \zeta) - \rho\cos\theta\left(L\frac{\partial F}{\partial x} + M\frac{\partial F}{\partial y} + N\frac{\partial F}{\partial z}\right) = 0. \quad \ldots(1)$$

Again, if $Q'(\xi', \eta', \zeta')$ be a point distant ρ' from P in the direction l, m, n, ρ' being small, the equation of the locus of Q' is

$$F(\xi' - l\rho', \eta' - m\rho', \zeta' - n\rho') = 0,$$

i.e. to the first approximation it is

$$F(\xi', \eta', \zeta') - \rho'\left(l\frac{\partial F}{\partial x} + m\frac{\partial F}{\partial y} + n\frac{\partial F}{\partial z}\right) = 0. \quad \ldots\ldots(2)$$

* T. Smith, *loc. cit.*

Digitized by Google

Equations (1) and (2) will represent the same surface if

$$\lambda\rho \cos\theta = \lambda\rho' \cos\theta',$$

i.e. if

$$\rho = \rho'.$$

Thus the locus of Q is the surface F displaced bodily, without rotation and without distortion, through a small distance ρ parallel to the fixed direction l, m, n.

§ 5. *The Optical Cosine Law.*

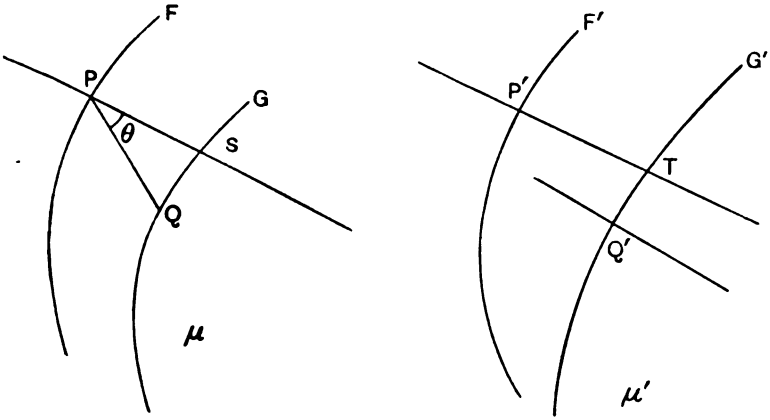


Fig. 1.

Let F be an incident wave surface in a medium of index μ and let F' be any position of the corresponding emergent wave in a medium of index μ'; let P and P' be corresponding points upon these two surfaces respectively, the reduced path between them being C_1, so that the normals at P and P' are corresponding 'rays.' Suppose the surface F moved, without rotation or distortion, to G through a small distance ρ parallel to some *fixed* direction inclined at an angle θ to the normal at P, Q being the new position of P; and let G' be a position of the emergent wave surface corresponding to G, the reduced path between them being C_2; let Q' be the point upon G' corresponding to Q. Then, if the normal at P' intersect G' in T and the normal at P intersect G in S, we have, in the notation of § 2,

$$V(QQ') = V(ST)$$
$$= V(PP') - \mu\rho\cos\theta + \mu' P'T;$$

i.e.,

$$\mu' P'T = \mu\rho\cos\theta + V(QQ') - V(PP'),$$

or

$$\mu' P'T = \mu\rho\cos\theta + C_2 - C_1, \qquad \ldots\ldots(1)$$

correct to small quantities of the first order.

Digitized by Google

Suppose now that for all rays we have

$$\cos\theta = \alpha\cos\theta' + \beta, \qquad \ldots\ldots(2)$$

α and β being constants and θ' the inclination of the emergent ray $P'T$ (or normal to the emergent wave surface) to some *fixed* direction in the image space, we may compare (2) with (1), which may be written

$$\cos\theta = \frac{1}{\mu\rho}(\mu' P'T - C_2 + C_1).$$

Since C_2 is arbitrary, we may write

$$C_1 - C_2 = \beta\mu\rho;$$

(2) is satisfied for all rays and therefore

$$\mu' P'T = \alpha\mu\rho\cos\theta'.$$

By the lemma of the preceding paragraph, therefore, the locus of T, i.e. the surface G', is merely a bodily displacement of the surface F', without rotation or distortion, through a small distance ρ' parallel to the fixed direction in the image space; and

$$\mu'\rho' = \alpha\mu\rho.$$

And inasmuch as a wave surface implies a corresponding caustic surface, we have the result:

If two systems of corresponding rays, touching two caustic surfaces, satisfy the condition $\cos\theta = \alpha\cos\theta' + \beta$, *$\theta$ and θ' being the inclinations of a ray to two fixed directions, the one in the image space and the other in the object space, and α and β being constants, then corresponding to a small displacement of either caustic surface, without rotation or distortion, parallel to the fixed direction associated with it, there will be a small displacement, without rotation or distortion, of the corresponding caustic surface parallel to the other fixed direction; and, if the displacements be respectively ρ and ρ' in media of indices μ and μ',* $\mu'\rho' = \alpha\mu\rho$.

This result constitutes the 'Optical Cosine Law'; and it is to be noticed that the fixed directions chosen are any whatever, and not necessarily optically corresponding ones. Moreover, the general caustic surface is considered so that the law contemplates the presence of any combination of the geometrical aberrations; and, further, it is not assumed that the optical system is in any way symmetrical, for the only principle employed is that of the 'reduced path.'

§ 6. It has been pointed out by the author of the Optical Cosine Law* that this law embodies a number of well-known results; we proceed to indicate some of them here.

* T. Smith, *loc. cit.*

Digitized by Google

(i) Let us consider a single refracting, or reflecting, surface as the optical system, and let one caustic degenerate to a point upon this surface; then the second caustic degenerates to the same point, and if the two fixed directions, chosen in accordance with the Cosine Law, coincide with any one tangent to the optical surface, the law of refraction, or of reflection, follows at once.

(ii) Consider now the general symmetrical optical system and assume two conjugate axial points, free from spherical aberration, so that the caustic surfaces degenerate again into these points: take parallel lines through them perpendicular to the axis of symmetry as the fixed directions. Then the condition associated with the Cosine Law is

$$\sin\theta = \alpha \sin\theta' + \beta,$$

θ and θ' being the inclinations of a ray with the axis of the system: and, since a ray coincides with the axis, i.e. $\theta = \theta' = 0$ must satisfy the condition, it follows that $\beta = 0$. Moreover, α being the factor of proportionality, and m the reduced magnification,

$$\alpha = m,$$

so that

$$\sin\theta = m \sin\theta',$$

i.e. the sine-condition. A slight movement of one point perpendicular to the axis of the system implies a slight parallel movement of the corresponding caustic surface, which remains a point, i.e. implies the absence of circular coma of all orders. But it is to be noticed that the Cosine Law is valid even in the presence of spherical aberration, and in this case also it will imply the absence of circular coma of all orders.

(iii) The argument of (ii) may be repeated, except that the axis itself may be taken as defining the two fixed directions; then with a similar notation, we have

$$\cos\theta = \alpha \cos\theta' + \beta.$$

Again $\theta = \theta' = 0$ defines a ray, coincident with the axis, leading to the relation

$$1 = \alpha + \beta,$$

so that

$$\sin(\theta/2) = \sqrt{\alpha}\,.\,\sin(\theta'/2).$$

Here, however, by Maxwell's elongation formula, α being again the factor of proportionality of corresponding small displacements

$$\alpha = m^2;$$

whence we have

$$\sin(\theta/2) = m \sin(\theta'/2),$$

i.e. Herschel's condition. Under this condition, therefore, the caustic surfaces, being initially points, remain points, i.e. the spherical aberration is stationary at zero value for small axial displacements.

Digitized by Google

(iv) Consider now the conditions implied by perfect definition over the conjugate planes associated with reduced magnification m: here again the caustic surfaces will degenerate into points and we may take as our fixed directions any two lines lying in the normal planes considered, so that, as in (ii), $\alpha = m$. In the notation therefore of § 2, L, M, N and L', M', N' being the direction cosines of a ray referred to parallel axes, the Cosine Law leads to the result

$$M - mM' = \beta,$$

and β is a constant depending upon the object point, or image point, chosen, i.e. we have the cosine-condition

$$M - mM' = \text{constant};$$

and similarly

$$N - mN' = \text{constant}.$$

These results are a generalisation of the sine-condition, applicable to non-axial points; they are the 'cosine conditions,' applicable to the outer parts of the field.

Digitized by Google

The effect of the orbital velocity of the electrons in heavy atoms on their stopping of α-particles. By L. H. THOMAS, B.A., Trinity College.

[*Read* 31 January, *received* 11 February 1927.]

Henderson* proposed a theory of the stopping of swift α-particles by matter. He treated the electrons in the atoms as free and at rest and ignored all collisions of the α-particle with them except those in which the electron would on that assumption gain sufficient energy to leave the atom. Fowler† has shown that Henderson's theory gives stopping-powers only of about 60 % of those observed. Fowler‡ also made a calculation of the stopping-power of hydrogen by combining the effects of close collisions treating the electron as free and slight collisions as perturbations of the electron's motion in a circular orbit. He obtained much better agreement. This method is, of course, the natural extension of Bohr's§ original theory to that model.

The following paper contains the results of some calculations of the stopping-powers of neon, argon, krypton and xenon. The electrons are supposed to be moving in orbits as calculated by Hartree|| to fit X-ray and optical data for sodium, potassium, rubidium and caesium. Following Henderson the electrons are treated as free and collisions are ignored unless the transfer of energy is sufficient to remove the electron from the atom, but the speed of the electron in its orbit is taken into account. Slighter collisions are neglected because the calculation of the perturbations involved would be very long. Their inclusion would probably not greatly increase the results.

The stopping-powers so calculated are compared with Gurney's¶ experimental values with which they are in fair general agreement. It is not claimed that these calculations have much importance; they merely show that a correct quantum theory calculation has not to explain any large numerical discrepancy.

1. *The formulae used for calculating the stopping-power.*

Suppose that an α-particle of charge E and mass M moves with velocity V through a cloud of N electrons per unit volume of

* Henderson, *Phil. Mag.*, vol. XLIV (1922), p. 680.

† Fowler, *Proc. Camb. Phil. Soc.*, vol. XXI (1923), p. 521.

‡ *Ibid.*, vol. XXII (1925), p. 792.

§ Bohr, *Phil. Mag.*, vol. XXV (1913), p. 10, vol. XXX (1915), p. 581.

|| Hartree, *Proc. Camb. Phil. Soc.*, vol. XXI (1923), p. 625. I am indebted to Mr Hartree for his kindness in giving me unpublished numerical values for these fields and orbits.

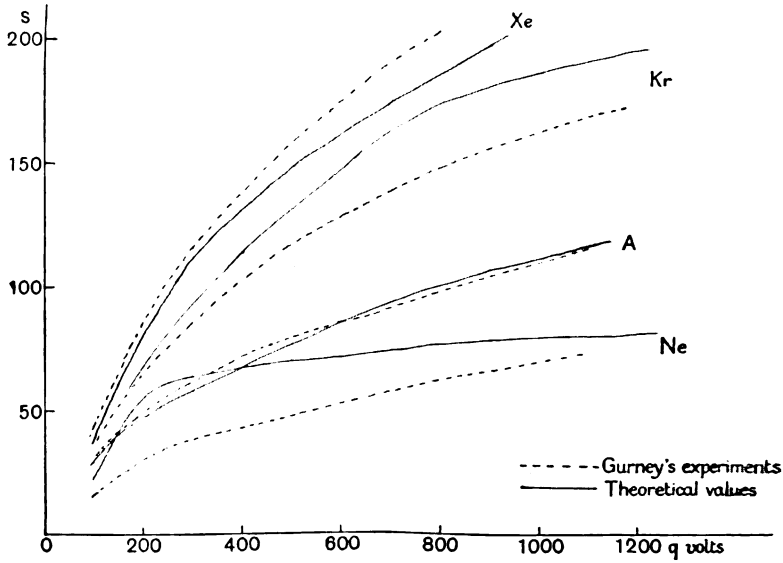
¶ Gurney, *Proc. Roy. Soc.* A. 107 (1925), pp. 332, 340.

Digitized by Google

charge e and mass m moving with velocity v in equal numbers in all directions. The chance of a collision in time dt in which r makes angle in range $d\chi$ about χ with V, the perpendicular distance of the electron from the line of relative motion of the particle lies in range dp about p, and the angle between the plane of the relative orbit and the (vV)-plane lies in range $d\epsilon$, is

$$N \tfrac{1}{2} d\cos\chi \, p dp d\epsilon u dt,$$

where u is the relative velocity before the collision*. By changing the variables and integrating over the relevant domain it follows that the chance of a collision in time dt in which the transference



of energy from the particle to the electron lies in range dq about q and of momentum in range ds about s is

$$\frac{4\pi N e^2 E'^2}{vVs^4} dt ds dq \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1),$$

if such a collision is possible at all. If V', v' are the velocities of the particle and electron after the collision, so that

$$q = \tfrac{1}{2} m (v'^2 - v^2) = \tfrac{1}{2} M (V^2 - V'^2) > 0,$$

the conditions that such a collision is possible are

$$m(v' - v) < s < m(v' + v),$$
$$M(V - V') < s < M(V + V').$$

* Cf. Jeans, *The Dynamical Theory of Gases*, p. 209.

Digitized by Google

Now suppose $M >> m$, so that the conditions are practically

$$m(v' - v) < s < m(v' + v),$$
$$\frac{m(v'^2 - v^2)}{2V} < s.$$

Integrating (1) over possible values of s, the chance of a collision in time dt in which the transference of energy is between q and $q + dq$ is

$$\left.\begin{array}{lll} \frac{4}{3}\frac{\pi Ne^2E^2dt}{m^3vV}\frac{(v+v')^3-(v'-v)^3}{(v'^2-v^2)^3}dq & \text{for} & \frac{v'^2-v^2}{2V} < v'-v \\ \frac{4}{3}\frac{\pi Ne^2E^2dt}{m^3vV}\frac{(2V)^3-(v'-v)^3}{(v'^2-v^2)^3}dq & \text{for} & v'-v < \frac{v'^2-v^2}{2V} < v'+v \\ \text{and} \quad 0 & \text{for} & v'+v < \frac{v'^2-v^2}{2V} \end{array}\right\} \quad \ldots\ldots\ldots\ldots(2).$$

The mean loss of energy by the particle in those collisions for which $q > j$ is found by integrating the product of q and (2), writing for brevity

$$k = \tfrac{1}{2}mv^2,$$
$$w = \tfrac{1}{2}mV^2,$$

to be $$\frac{2\pi Ne^2E^2dt}{mV}\left[\log\frac{4(w-k)}{j} + \frac{4}{3}\frac{k}{j}\right]^*$$

for $$4\{w - \sqrt{(wk)}\} > j,$$

$$\frac{2\pi Ne^2E^2dt}{mV}\left[\frac{4}{3}\frac{w}{j}\left(\frac{w}{k}\right)^{\frac{1}{2}} + \log\left\{1+\left(\frac{w}{k}\right)^{\frac{1}{2}}\right\} - \left(\frac{w}{k}\right)^{\frac{1}{2}}\right.$$
$$\left. + \log 2 - \log\left\{1+\left(1+\frac{j}{k}\right)^{\frac{1}{2}}\right\} + \frac{1}{3} - \frac{1}{3}\left(2\frac{k}{j}-1\right)\left\{\left(1+\frac{j}{k}\right)^{\frac{1}{2}} - 1\right\}\right]$$

for $$4\{w + \sqrt{(wk)}\} > j > 4\{w - \sqrt{(wk)}\},$$

and 0 for $$j > 4\{w + \sqrt{(wk)}\} \ldots\ldots\ldots\ldots\ldots\ldots(3).$$

2. *The numerical results.*

It is convenient to compare with experiment the coefficient in the V^4-law, or rather the stopping-power S as defined by Fowler, which is given by

$$-\frac{dV^4}{dl} = \frac{8\pi Ne^2E^2}{mM}S,$$

* The extra term $4k/3j$ in this expression when the velocity of the electron is taken into account depends on limiting the collisions by a restriction on q. In Fowler's calculation (*loc. cit.*) for hydrogen the close collisions are in effect separated from the others at a value of p. The only alteration that taking the velocity in close collisions into account makes in his result is to replace V^3 in the numerator of the argument of the logarithm in his formula ((1), p. 794, *loc. cit.*) by $V^2(V^2 - v^2)^{\frac{1}{2}}$.

Digitized by Google

where V is the velocity of the particle at distance l along its path, and N is the number of atoms per unit volume of the matter traversed.

In accordance with the assumptions stated above values of S were calculated to ten-inch slide-rule accuracy by summing over the electrons in an atom the time averages with respect to k over their respective orbits of the expressions in square brackets in (3), j being replaced by the energy required for ionisation. The results are given in the following table and are compared with Gurney's experiments in the accompanying graph. 'q' is the energy of the α-particle in 'electron-volts,' i.e. the number of volts through which an electron must fall to attain the velocity of the α-particle.

For argon and xenon there is very good agreement. For neon the theoretical results are too large at small speeds. However, Hartree's field for sodium used in the calculation is itself less accurate than the other fields (for K, Rb and Cs) and would be expected to differ more from that for neon than the others do from those for the corresponding inert gases. The theoretical values for krypton are too large at high velocities, for which, however, the bulk of the stopping is done by the ten electrons in 3_3 orbits. Near these orbits Hartree's field for rubidium is nearly an inverse cube field, and a small difference between the fields for rubidium and krypton might mean a large difference in the velocity of an electron in a 3_3 orbit and so a large difference in the calculated stopping-power.

TABLE. *Calculated values of S.*

q (volts)...	100	200	300	400	500	600
Ne	23·1	55·2	63·5	67·9	70·8	72·2
A	28·3		57·7		77·0	
Kr		66·6		114·2		147·7
Xe	44·0		112·1		148·9	

q (volts)...	700	800	900	1000	1100	1200
Ne	75·1	77·7	78·5	79·9	81·3	82·2
A	94·7		107·5		116·4	
Kr		175·0		186·9		196·0
Xe	175·5		197·9		217·0	

Digitized by Google

The Absolute Intensities of the γ-Rays of Radium B and Radium C. By C. D. ELLIS, Ph.D., Trinity College, and W. A. WOOSTER, B.A., Peterhouse.

[*Received* 18 February, *read* 14 March 1927.]

Introduction.

Considerable advances have been made in recent years in our knowledge of the γ-ray emission of radioactive bodies. It has been established quite definitely that series of definite frequencies are emitted, forming a characteristic line spectrum of the nucleus, and methods of measuring the frequencies have been devised so that at present the γ-ray spectra of most of the radioactive bodies are well known.

Recently the centre of interest has moved rather away from the numerical values of the frequencies, since so far little progress has been made in finding any series formula, and more attention is being directed to the intensities. Estimates of the intensities of the γ-rays of radium B and C have been made independently by Ahmad* and Kohlrausch†, and one interesting result is that these authors find difficulty in accounting for all the observed effects of absorption and scattering, using the distribution of intensities that they deduce, so that among other possibilities they are led to consider the emission of continuous γ-radiation as well as the characteristic line spectrum. If this were to be true it would have important and fundamental bearing on our picture of the β-ray disintegration, but it is our opinion that it is too early to say anything with certainty. It is with this point in view that we wish to discuss an alternative method of estimating the intensities which brings out clearly the exact points of uncertainty in our present knowledge, and indicates how they can be removed. We think it is undesirable to attempt to settle an important question like the existence or non-existence of the continuous γ-ray spectrum on insufficient evidence, especially as there is a reasonable hope that experiment can soon give an unambiguous answer. However, until this is carried out, we must depend on estimates of the intensities, and a further object of this paper is to describe an estimate of the intensities of the γ-rays, which is both more accurate and more detailed than any previously given.

* Ahmad, *Proc. Roy. Soc.* 109, p. 207. 1925.
† Kohlrausch, *Phys. Zeit.* 28, p. 1. 1927.

Digitized by Google

General Discussion of Method.

The method that was used by Ahmad* was to make measurements of the absorption of the complex radiation from radium B and C after it had passed through various filters, and then, assuming laws for the absorption of the individual constituents of the beam, he deduced what must be the relative intensities of these constituents in order to give the observed results. In a somewhat similar manner Kohlrausch†, making use of Thibaud's‡ estimated intensities of the groups of electrons ejected by the γ-rays from ordinary matter, deduces values for the relative intensities of the different frequencies, and then checks his results by comparison with known absorption and scattering experiments. We will not refer to the results of these two authors in more detail, since at this stage it is sufficient to point out that both depend on the laws of absorption and scattering of high frequency radiation. There is very little of a definite nature known about these laws, the best that can be done is to extrapolate the X-ray results, and it is therefore of great interest to form an estimate of the γ-ray intensities by a method which does not depend on these laws.

The general principle of our method is as follows. The homogeneous groups in the β-ray spectra are due to the conversion of the γ-rays, and the intensities of these groups depend not only on the intensities of the γ-rays but also on the extent to which they are converted, that is the magnitude of the internal conversion coefficient. Conversely if the value of the internal conversion coefficient is known for all the different γ-ray frequencies, then we have only to divide the intensities of the β-ray groups by these coefficients in order to obtain the intensities. If we are only concerned with relative intensities of the γ-rays, then it would be sufficient to use relative intensities of the β-ray groups and to know the rate of the variation with frequency of the internal conversion coefficient without troubling about the actual values. However the absolute intensities of the γ-rays are far more interesting, and by these we mean the actual number of quanta of the different frequencies emitted in, say, a thousand disintegrations, and to obtain these we require to know the number of electrons in the β-ray groups from a known number of disintegrations and also the actual values of the internal conversion coefficient. We have recently published values for what may be called the absolute intensities of the β-ray groups, and so the problem that remains to be settled is to find values of the internal conversion coefficient. We hope soon to measure these experimentally, but in this

* *Loc. cit.* † *Loc. cit.*

‡ Thibaud, *Journ. de Phys.* 34, p. 807. 1925.

Digitized by Google

paper we deduce them by comparison with measurements of the total energy of the γ-rays, and further with the total number of quanta emitted*. It is easy to see in a general way how this is possible. Supposing it be assumed that the internal conversion coefficient varies only slowly with the frequency, then the intensities of the γ-rays will not differ greatly from the intensities of the β-ray groups, and since we know that the low energy β-ray groups are most intense, this will result in a concentration of the γ-ray energy in the low frequencies. Then to account for the large amount of energy emitted as γ-rays, we shall require a great number of quanta of these low frequencies, more in fact than are found by direct counting experiments. If, on the other hand, we assume a very rapid decrease of the internal conversion coefficient, then we find we can account for the total energy with fewer quanta than are actually found.

In this way we are able to choose a definite rate of variation with frequency, and definite absolute values for the internal conversion coefficient, and then combining this result with the intensities of the β-ray groups, we obtain the intensities of the γ-rays.

Since energy and counting measurements have been made for both radium B and C, and since it is reasonable to assume the same rate of variation of the internal conversion coefficient for both bodies, it is possible to obtain a check on our results. It is strong evidence in support of our method that the separate calculations all lead to the same results.

Theory of the Method of Estimation.

Consider a single radioactive body which emits n homogeneous radiations whose frequencies (and therefore also the energies of individual quanta), measured in volts, are

$$V_1, V_2, V_3, \ldots V_n.$$

All of these frequencies are not emitted at each disintegration, on the contrary we must associate with each frequency a probability p, so that out of a very large number N of disintegrations there will be emitted Np_1 quanta of frequency V_1, Np_2 of frequency V_2, and so on. The quantities p_1, p_2, etc. will be called the absolute intensities of the γ-rays. We shall use these intensities p to denote the probabilities of emission of the γ-rays from the nucleus; they will not represent, except approximately, the intensities of the γ-rays that finally emerge from the atom because some of the γ-rays are converted in the parent atom, giving rise to the groups

* It will be seen later that, while advantageous, it is not necessary to depend on the counting measurements, but the method is most easily explained by referring to them.

Digitized by Google

of the natural β-ray spectrum. We must take into consideration that the probability α of this internal conversion varies with the frequency and the level, K, L, M, etc., in which the conversion occurs. Using an obvious notation we write

$$ {}_K\alpha_1,\ {}_{L_I}\alpha_1,\ \cdots\ {}_K\alpha_2\ \cdots $$

for these coefficients.

We can now write down expressions for the total γ-ray energy emitted from the nucleus per disintegrating atom

$$ E_{\text{nucleus}} = \sum_1^n p_r V_r. \qquad \ldots\ldots(1) $$

The amount of γ-ray energy not converted in the K level is $pV(1 - {}_K\alpha)$, and continuing this we find for the total γ-ray energy finally emerging from the atom

$$ E = \sum_1^n p_r V_r (1 - {}_K\alpha_r)(1 - {}_{L_I}\alpha_r) \ldots (1 - {}_M\alpha_r). \quad \ldots\ldots(2) $$

It is this last quantity which is determined by experiment, and with which we shall have to deal. It is frequently referred to as the heating effect of the γ-rays.

It is known that the coefficients of internal conversion are small, that due to the K level (${}_K\alpha_r$) being in general less than $\frac{1}{10}$, whereas the L and M coefficients are much smaller. We can therefore write

$$ E = \sum_1^n p_r V_r [1 - ({}_K\alpha_r + {}_{L_I}\alpha_r + \ldots {}_M\alpha_r \ldots)] \quad \ldots\ldots(3) $$

with sufficient accuracy, and then further, making use of a result that we have established recently*, that to a sufficient approximation, independent of the frequency

$$ {}_{L_I}\alpha + {}_M\alpha = 0{\cdot}22\,{}_K\alpha, $$

we have finally

$$ E = \sum_1^n p_r V_r (1 - 1{\cdot}22\,{}_K\alpha_r). \qquad \ldots\ldots(4) $$

In a similar manner the number N of quanta emitted clear of the atom is given by

$$ N = \sum_1^n p_r (1 - 1{\cdot}22\,{}_K\alpha_r). \qquad \ldots\ldots(5) $$

This number has been measured by counting the separate γ-ray impulses with an electrical counter.

These two equations, (4) and (5), are clearly not sufficient by themselves to determine the probabilities of emission p, but this

* Ellis and Wooster, *Proc. Roy. Soc.*, March, 1927.

Digitized by Google

can be achieved by considering the intensities of the β-ray groups. If we denote by $n\theta_1$, $n\theta_2$, $n\theta_3$, ..., the number of electrons in the different groups from a large number n of disintegrating atoms due to the conversion of the γ-rays of frequency V_1, V_2, V_3, ... in the K level, then we may write

$$\theta_1 = p_{1K}\alpha_1, \quad \theta_2 = p_{2K}\alpha_2, \ldots. \qquad \ldots\ldots(6)$$

Values for the quantities θ, the intensities of the β-ray groups, have been obtained in the case of radium B and C*.

It is now only a question of manipulating equations (4), (5) and (6) so as to obtain the p's, the γ-ray intensities. Instead of doing this directly, we fix our attention first on the internal conversion coefficient α. We assume that this varies smoothly with frequency in a manner represented by $f(V)$. It is convenient to write g for the value at a standard frequency (actually $V = 3{\cdot}54 \times 10^5$ volts, the strongest radium B γ-ray) so that, for this value of V, $f(V)$ is equal to unity, then at any other frequency the internal conversion coefficient in the K level will be $gf(V)$.

The intensities of the β-ray groups θ may now be written

$$\theta_r = gf(V_r)\,p_r. \quad (r = 1, 2, 3 \ldots n) \qquad (7)$$

Returning to the equations (4) and (5) we substitute $\theta_1, \theta_2, \ldots$ for the products $p_{1K}\alpha_1 \ldots$, according to equation (6), and further $\theta_1/gf(V_1)$ for p_1, according to equation (7).

This gives the two final equations

$$E + 1{\cdot}22 \sum_1^n \theta_r V_r = \frac{1}{g} \sum_1^n \frac{\theta_r V_r}{f(V_r)}, \qquad \ldots\ldots(8)$$

$$N + 1{\cdot}22 \sum_1^n \theta_r = \frac{1}{g} \sum_1^n \frac{\theta_r}{f(V_r)}. \qquad \ldots\ldots(9)$$

Now in these equations experimental values can be assigned to the frequencies of the γ-rays (V), the intensities of the β-ray groups (θ), the total energy of the emitted γ-rays (E), and the total number of emitted γ-rays N. The whole problem is to make such a choice of the value of g, and the form of the function $f(V)$, that these two equations may both be satisfied. Once this is carried out, equation (7) gives the required intensities of the γ-rays.

It will always be possible to find values for g and $f(V)$ so that equations (8) and (9) are satisfied within the experimental error, but a very important check on the correctness of our point of view may be obtained by considering different radioactive bodies. The value of g, the conversion coefficient for a fixed frequency, might be expected to show a variation with atomic number, but since we are only going to apply these calculations to radium B and

* Ellis and Wooster, *loc. cit.*

Digitized by Google

radium C (atomic numbers when the γ-rays are emitted 83 and 84), there will be only a change of 1·2 per cent. in the atomic number. The external absorption coefficient, with which we might compare the internal conversion coefficient, would only change by 5 per cent. in this case, so that it will be quite sufficient, in view of the other inaccuracies, if we anticipate a small increase in g for Ra C and write

$$g_{\text{Ra C}} = 1{\cdot}03\, g_{\text{Ra B}}.$$

The function $f(V)$ depending only on the frequencies of the γ-rays should be the same for all radioactive bodies.

Thus finally we obtain a check on our hypotheses by demanding that we account for the energies and numbers of the γ-rays from both radium B and radium C with the same function $f(V)$ and values of g differing only by about 3 per cent.

Estimation of the Internal Conversion Coefficient and its Variation with Frequency from measurements of the total energy and total number of emitted γ-rays.

We will first describe shortly the experimental material on which we are basing our estimate. This all refers to radium B and C, and consists of measurements of the total number of γ-rays, the total energy of the γ-rays, the frequencies of the γ-rays and the intensities of the β-ray groups.

The total number of γ-ray impulses emitted by a known quantity of radium B and radium C has been measured by Kovarik* using an electrical counter. Two series of measurements were carried out, in one the radioactive material was only covered by thin glass, and all the γ-rays of radium B and C, except the softest of radium B, were counted. In the second case the source was covered by 15 mm. of lead, so that only the hard γ-rays of radium C were able to penetrate. This latter measurement, when corrected for absorption, will give an approximation to the number of hard γ-rays emitted by radium C. These results, giving $N' = 2{\cdot}14$ γ-impulses from 1 atom of Ra B and 1 of Ra C, and further $n = 1{\cdot}63$ γ-impulses from 1 atom of radium C, should be correct to about 5 %†.

The present authors‡ have measured, by means of its heating effect, the total energy emitted in γ-ray form during the disintegration of one atom of radium B and one of radium C, and have found the value

$$E' = 18{\cdot}4 \times 10^5 \text{ volts}.$$

* Kovarik, *Phys. Rev.* [5], p. 559. 1924. See also footnote to equations (10) and (11) later.

† Assuming $3{\cdot}4 \times 10^7$ atoms disintegrate per sec. from 1 mg. This figure is used throughout this paper; any variation in this figure will only affect the value of "g" and not $f(V)$.

‡ Ellis and Wooster, *Phil. Mag.* [6] L, p. 521 (1925).

Digitized by Google

This figure should also be correct to about 5 %.

The division of this energy between radium B and radium C has been estimated by various means, either using the relative absorption coefficients, or the relative numbers of ions formed in air. Combining the values of J. Szmidt*, Moseley and Robinson†, Moseley and Makower‡, Ellis and Wooster§, we think

$$\epsilon' = \text{Energy Ra B/Energy Ra C} = 0{\cdot}085$$

is the best value, and is probably correct to 10 %.

It would be possible to include all the known γ-rays of radium B and C, but we think it sufficient in this preliminary estimate to include only seven from each. The frequencies measured in volts are shown in Table I, and also in the third column the corresponding values of θ, the intensities of the β-ray groups due to conversion of the corresponding γ-ray in the *K* level. These intensities give directly the number of electrons per disintegrating atom. The values of the frequencies are well known and probably accurate to one part in 300; the intensities of the β-ray groups are taken from the authors' recent measurements, and may be in error from 5 to 10 %. We do not include the lowest frequency γ-ray of radium B because, on the one hand, it contributes little to the total energy, and, on the other hand, it was probably not measured by Kovarik.

TABLE I. *Data of γ-ray Frequencies and Intensities of β-ray groups used in Calculations.*

Number of γ-Ray		Frequency in Volts V	Intensity of β-ray group due to *K*-conversion of γ-ray θ
Radium B	1	$2{\cdot}43 \times 10^5$	0·0314
	2	2·60	0·0038
	3	2·76	0·0019
	4	2·97	0·0435
	5	3·54	0·0550
	6	4·71	0·0008
	7	4·82	0·0008
Radium C	8	6·12	0·00396
	9	9·41	0·00047
	10	11·30	0·00107
	11	12·48	0·00044
	12	14·26	0·00223
	13	17·78	0·00046
	14	22·19	0·00011

* *Phil. Mag.* [6] XXVIII, p. 527 (1914).
‡ *Ib.* [6] XXIII, p. 302 (1912).
† *Ib.* [6] XXV, p. 312 (1913).
§ *Ib.* [6] L, p. 521 (1925).

Digitized by Google

We have now collected all the necessary material in order to estimate the value "g" of the internal conversion coefficient for our standard frequency and its variation $f(V)$ with frequency V, and we may now insert the numerical values in the equations of the last paragraph.

Applying equation (9) to the total number of γ-rays from radium B and C*

$$\sum_{1}^{7} \frac{\theta_r}{f(V_r)} + \frac{1}{1{\cdot}03} \sum_{8}^{14} \frac{\theta_r}{f(V_r)} = g\left[N' + 1{\cdot}22 \sum_{1}^{14} \theta_r\right] = 2{\cdot}32g. \quad \ldots(10)$$

Applying the same equation for the number of γ-rays from radium C alone

$$\sum_{8}^{14} \frac{\theta_r}{f(V_r)} = 1{\cdot}03g\left[n' + 1{\cdot}22 \sum_{8}^{14} \theta_r\right] = 1{\cdot}79g. \quad \ldots\ldots(11)$$

Applying equation (8) to the total energy of the γ-rays from radium B and C

$$\sum_{1}^{7} \frac{\theta_r V_r}{f(V_r)} + \frac{1}{1{\cdot}03} \sum_{8}^{14} \frac{\theta_r V_r}{f(V_r)} = g\left[E' + 1{\cdot}22 \sum_{1}^{14} \theta_r V_r\right] = 19{\cdot}1g. \quad (12)$$

Using equation (8) to give the ratio of the energy of the γ-rays from radium B and C

$$\frac{\sum_{1}^{7} \frac{\theta_r V_r}{f(V_r)}}{\sum_{8}^{14} \frac{\theta_r V_r}{f(V_r)}} = \frac{1}{1{\cdot}03} \frac{\epsilon' + \frac{1{\cdot}22\,(1+\epsilon')}{E'} \sum_{1}^{7} \theta_r V_r}{1 + \frac{1{\cdot}22\,(1+\epsilon')}{E'} \sum_{8}^{14} \theta_r V_r} = 0{\cdot}11\,(2). \quad \ldots(13)$$

In these four equations the various values of θ and V are of course known and may be taken from Table I. The entire problem consists in choosing a value for g and a form of the function $f(V)$ valid between $V = 2{\cdot}0 \times 10^5$ and $25{\cdot}0 \times 10^5$ with value unity at $V = 3{\cdot}54 \times 10^5$, so that these four equations are satisfied.

Perhaps the easiest way of finding a solution is to choose arbitrarily some form for $f(V)$; this may be at once checked on equation (4), and also used to find values for "g" from each of the equations (10), (11) and (12). If it is possible to find a $f(V)$ satisfying (13) and leading to the same values of "g" in equations (10), (11) and (12), then with some confidence this may be used to deduce the intensities of the γ-rays from the intensities of the β-ray groups†. This point will be most easily understood by

* The frequent appearance of the factor 1·03 in these equations is due to the slightly larger values of the internal conversion coefficients for radium C. It will be noted that the above equations give simply the energies and numbers of the γ-rays emitted from the nucleus before conversion takes place.

† It is generally recognised that the counting of γ-rays is an extremely difficult experiment, and so our results based on these measurements might appear to be open to doubt. It has been suggested that Kovarik's figure might be too large, owing

Digitized by Google

reference to the next table which shows the result of trying a variation of the internal conversion coefficient with the inverse 2·5, 2·6, 2·65, 2·7 and 2·75 power of the frequency.

TABLE II. *Correlation of γ-Ray Heating and γ-Ray Counting Measurements by means of different rates of variation of the Internal Conversion Coefficient.*

Internal Conversion Coefficient varying as	$V^{-2.5}$	$V^{-2.6}$	$V^{-2.65}$	$V^{-2.7}$	$V^{-2.75}$
Value of "g" deduced from					
1. Number of γ-rays emitted by radium B and C	11·0	11·8	12·4	13·1	14·0
2. Number of γ-rays emitted by radium C ...	8·9	10·2	11·0	11·8	12·8
3. Energy of γ-rays emitted by radium B and C	7·9	9·7	12·6	16·9	24·3
Ratio $\frac{\text{Energy γ-rays of radium B}}{\text{Energy γ-rays of radium C}}$	Calculated				
Exper. 0·11	0·27	0·21	0·16	0·11	0·045

The criterion of a correct choice of the rate of variation would be that the values of g shown by the first three figures in any vertical column should be identical, and further the figure in the last row should be 0·11. It will be seen that a rate of variation between $V^{-2.6}$ and $V^{-2.65}$ would satisfy the first three criteria but that the ratio of the energy of the radium B γ-rays to those of radium C favours definitely the larger value of the exponent. It would be no doubt possible to find other methods of variation of the internal conversion coefficient which would give closer agreement, but although they might differ greatly in the algebraic form used to represent them they would not differ much in the numerical values over the range considered, and as a basis for a provisional estimate we shall take the internal conversion coefficient to vary as the inverse 2·65 power of the frequency and to have the value 0·12 at a frequency $3{\cdot}54 \times 10^5$ volts*.

to a high speed photo-electron giving rise, by collision, to another fast electron. We do not share this view, and consider Kovarik's experiments may be used with safety, but it is most important to note that, in point of fact, we could dispense with these measurements altogether. Using equations (12) and (13) we could deduce values of "g" and $f(V)$, which we should then have to justify on general grounds. Then with these values we could use equations (10) and (11) to calculate the number of γ-rays emitted.

* Were we to leave out the counting measurements we should decide on a variation with the 2·7 power and a value for g of 0·17. This would put still more energy in the high frequencies.

Digitized by Google

It is a very striking point that these values are in excellent accord with what we should expect on general considerations. On the one hand, it has long been known that the value of the coefficient for the radium B γ-rays was of this general order of magnitude; on the other hand, Gurney deduced that for the strongest γ-ray of thorium B ($V = 2{\cdot}41 \times 10^5$) the conversion coefficient must be greater than 0·25 in order to account for the β-ray groups with not more than one quantum emitted at each disintegration.

Now as we have already stated it seems probable that the internal conversion coefficient will depend, apart from the frequency of the γ-ray, only on the electronic structure where the conversion takes place, and therefore should be the same for thorium B as for radium B.

If we extrapolate our value of 0·12 back to the thorium B frequency according to an inverse 2·65 power, we obtain a value 0·33 which it is very satisfactory to note is greater than the lower limit 0·25 assigned on general grounds above.

As regards the rate of variation according to the inverse 2·65 power of the frequency, we may note that this is another point in favour of the general conclusion that internal and external absorption are essentially similar. It is well known that the ordinary absorption varies approximately as the inverse third power of the frequency, and if anything slightly less rapidly. Recent theoretical investigations have in fact suggested a variation for absorption in the K level according to the inverse 2·5th power; but without entering into further detail, it is clear that to a first approximation the internal and external absorption coefficients vary in the same general manner.

Values of the Intensities of the γ-Rays.

The next table shows the intensities of the γ-rays obtained by dividing the intensities of the β-ray groups (see Table I) by the value of the internal conversion coefficient. For the reasons given in the last paragraph this has been taken to have the value 0·12 at a frequency $V = 3{\cdot}54 \times 10^5$ volts and to vary as the inverse 2·65 power of the frequency.

It should be noted that these are not only relative intensities but actually give the average number of quanta (less than unity) of the particular frequency emitted at each disintegration.

The most powerful γ-ray is one of 1·4 million volts from radium C, a quantum of which is emitted three times out of four disintegrations, but yet owing to the low value of the internal conversion coefficient at this frequency the resulting β-ray line is actually less intense than that resulting from the first γ-ray of radium B, one quantum of which is emitted in every ten disintegrations.

Digitized by Google

TABLE III. *Intensities of the γ-rays of Radium B and Radium C.*

Frequency measured in volts	Intensity Average number of quanta emitted per disintegration
RADIUM B	
$2{\cdot}43 \times 10^5$	0·095
2·60	0·014
2·76	0·008
2·97	0·228
3·54	0·458
4·71	0·014
4·82	0·015
RADIUM C	
6·12	0·140
9·41	0·051
11·30	0·193
12·48	0·104
14·26	0·748
17·78	0·274
22·19	0·119

These intensities are only preliminary estimates and it would be unwise to base any detailed deductions on them but, considered as an illustration, we think our present values bring out the great interest of the method in that it gives the absolute intensities. It promises to provide a far deeper insight into the mechanism of the γ-ray emission, since we can follow in detail the emission of the different quanta at each disintegration and further, if the evidence about stationary states in the nucleus be accepted as conclusive, we can follow the actual transitions occurring at the disintegration.

Previous estimates of the γ-ray intensities have only given relative values. For instance Ahmad* found that to account for the absorption and scattering of γ-rays, using extrapolated X-ray laws, it was necessary to assume relative intensities which made the first γ-ray of radium C six times stronger than the 1·4 million γ-ray. We find exactly the opposite, the 1·4 million γ-ray being five times stronger than the lower frequency γ-ray. This is a large and striking disagreement, and there seems little doubt that, making all possible allowances for uncertainty in data, the explanation lies in the inapplicability of the X-ray laws of absorption and scattering to these high frequency radiations. The necessary alteration will

* Ahmad, *loc. cit.*

Digitized by Google

be in the direction of a less rapid variation with wave-length of the photo-electric absorption, a result already suggested on theoretical grounds independently by Oppenheimer and Wentzel.

The recent careful analysis by Kohlrausch* of the absorption and scattering results lays down clearly the method to be followed and shows the valuable information that can be obtained if instead of attempting to deduce the intensities from the absorption results, the method is reversed, and the intensities, obtained for instance as in this paper, are used to predict the absorption and scattering that would be obtained on any particular laws.

As Kohlrausch points out, he had no reliable estimates of the γ-ray intensities from which to start, but to illustrate his method he made use of some visual estimates of β-ray photographs published by Thibaud. Kohlrausch's assumed relative intensities are not greatly different from those published here as regards general order of magnitude, although undoubtedly he has overestimated the intensities of the high frequency radiations. However, modifying his results to take this into account his conclusion that the absorption laws are in error is still valid. A development of this method, based on accurate data, promises to yield definite and precise information about the behaviour of high-frequency radiation.

Conclusion.

The object of this paper has been to show how the heating effect of the γ-rays, the total number of the quanta emitted, and lastly the intensities of the β-ray groups may all be correlated in terms of the absolute intensities of the γ-rays.

The chief point of the method is that the correlation of these experimental measurements depends only on the value and rate of variation with frequency of the internal conversion coefficient of the γ-rays, and once this has been ascertained the intensities of the γ-rays follow at once from the intensities of the β-ray groups. It is of great interest that the values of the internal conversion coefficient, which are found to be necessary in order to connect both the total energy and number of the γ-rays with the measurements of the intensities of the β-ray groups, are in entire agreement with previous estimates based on general considerations.

This provides considerable evidence in favour of the correctness of the calculations, and the values of the γ-ray intensities given in this paper represent in our opinion the best estimates that can be made at present.

The whole problem would be placed on a much firmer footing if independent measurements could be made of the rate of variation

* *Loc. cit.*

Digitized by Google

of the internal conversion coefficient, so that instead of having to deduce its values so as to be able to account for the total energy and number of the γ-rays, we could use the measured values of the conversion coefficient and ascertain whether the known monochromatic γ-rays were present in sufficient intensity to account for the total output of energy. This would appear to be the most promising way of finding out whether continuous γ-rays are emitted or not. If more γ-ray energy is given out than can be accounted for by the monochromatic γ-rays, then there is no alternative but to assume the existence of a continuous spectrum, but if on the contrary there is no surplus of energy, then, quite definitely, there can be no continuous γ-rays emitted.

We feel that it is very dangerous to deduce anything about the existence of continuous γ-rays from the non-agreement of theoretical expectations with the result of absorption experiments, since it is just as likely that our knowledge of the absorption of high-frequency radiation is at fault. It would appear much safer to decide the problem of the possible existence of the continuous spectrum of γ-rays by the method we propose, involving essentially a measurement of the rate of variation of the internal conversion coefficient.

We hope soon to obtain some information on this point, but even now it is permissible to point out, that since we have accounted for the total energy output with what is certainly a reasonable value for the conversion coefficient, the balance of evidence is already distinctly against the existence of a continuous γ-ray spectrum.

Digitized by Google

An Experiment on the Stability of Superposed Streams of Fluid. By Mr G. I. TAYLOR, Trinity College.

[*Read* 14 March 1927.]

It is well known that the calm which commonly comes on a cold clear night is due to the cooling effect of the ground on the lowest layers of the atmosphere. It frequently happens that the wind at a height of a few hundred feet does not die down at all at night; in fact observations taken at the top of the Eiffel tower show that there is frequently a slight increase in wind at that height. The cooling of the air near the ground makes the lower layers heavier than the higher layers and thus tends to prevent the formation of turbulence at the earth's surface. Since it is only through the medium of turbulence that the higher layers are able to drag the lowest layers over the ground the surface air stops moving.

The experiment here described was designed to illustrate the stabilising effect of a density distribution similar to that which occurs in the air near the ground at night. The apparatus, which is designed so that the experiment can be shown by means of a lantern, is illustrated in Fig. 1.

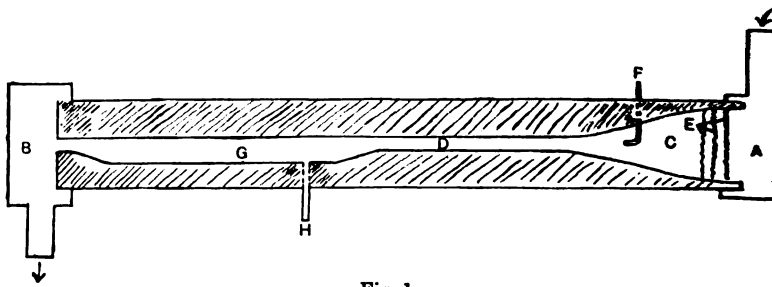


Fig. 1.

Two glass plates 6 cm. × 50 cm. are separated by two pieces of wood 6 mm. thick, the four components being stuck together with wax. The channel thus formed is used to convey water from the entrance chamber *A* (Fig. 1) to the exit chamber *B*. The entrance to this channel is shaped so that the converging stream of water passes smoothly from the entrance chamber *A* to the parallel part *D* of the channel. In order to damp down any large eddies which might be formed by the fluid on entering the converging space *C* three pieces of brass gauze *E* are fitted across the channel close to the entrance. Into the converging part is fitted a small capillary tube *F* bent at right angles at the lower end. Eocene is made to flow through this, and at low speeds when the flow is stream line

Digitized by Google

the colouring matter lies in a thin thread as in Osborne Reynolds' well-known experiment, down the middle of the channel. If the speed is increased the appearance of this thread of colouring matter is unaltered till turbulence sets in, when it spreads out, the colouring matter being diffused across the tube nearly to the walls.

After passing through 14 cm. of uniform cross-section the stream passes over a section containing brine. The coloured brine is led in through the pipe H to the bottom of a shallow trough G (Fig. 1), which is formed by a local increase in depth of the channel.

At low speeds of flow the surface of separation between the brine and the water is quite distinct, but there is a tendency for the stream to carry the brine out with it towards the exit, and this is partially counteracted by tilting the apparatus slightly so that the surface of separation is parallel to the centre line of the channel. Some of the brine is continually being removed by viscous drag however, and this is replaced by a slow continuous flow of brine delivered at H.

At slow speeds the flow is stream line and the surface of separation at the top of the brine is quite distinct. When the speed increases so that the upper part becomes turbulent the surface of separation still remains distinct, till, at a certain speed, instability sets in and all the brine, except a small quantity close to the place where the upper stream first comes into contact with it, is suddenly carried out of the channel.

If the same experiment is tried, using coloured water instead of brine, the appearance at slow speeds is similar to that just described. Soon after turbulence sets in, however, the lower liquid begins to mix with the upper one. The surface of separation remains distinct close to the place where the upper stream first encounters the lower coloured fluid, but mixture begins at a short distance down stream and continues with greater and greater intensity till the exit is reached.

The experiment may perhaps be regarded as an illustration of the way in which a cold and frequently misty patch of air lying in a hollow is suddenly blown away by a gradually rising wind in the higher layers.

Digitized by Google

The Stopping Power of Hydrogen Atoms for α-Particles according to the New Quantum Theory. By J. A. GAUNT, B.A., Trinity College. (Communicated by Mr R. H. FOWLER.)

[*Received* 2 March, *read* 14 March 1927.]

1. *Introduction.*

An α-particle passing through a gas loses velocity because it gives up energy to the gaseous atoms. If we can calculate the average energy transferred to an atom, the stopping power of the gas follows immediately. The earlier theories, by Thomson* and Darwin†, took account of only the close collisions, in which the α-particle actually passes through the atom. Bohr‡, however, in 1913, took into account the transfers of energy to atoms at distances from the track considerably larger than atomic dimensions. His calculation was purely classical, and dealt with an atomic model in which the electrons were capable of simple harmonic motions. The result was in very satisfactory agreement with experiment. For atoms containing each a single electron, with mass μ, charge ϵ, and natural period ω, Bohr's formula is

$$-\frac{dT}{dx} = \frac{4\pi N \mathrm{E}^2\epsilon^2}{\mu v^2} \log \frac{\gamma\mu v^2}{2\pi \mathrm{E}\epsilon\omega} \quad \ldots\ldots\ldots\ldots\ldots\ldots (1{\cdot}0),$$

where $-dT/dx$ is the rate of loss of energy by an α-particle whose charge is E, and velocity v; N is the number of atoms per c.c.; and $\gamma = 1{\cdot}123$.

The calculation was repeated by Fowler§ in 1925, with the usual planetary model of the hydrogen atom, instead of a harmonic oscillator. In its undisturbed state, the electron rotated in a circular orbit (the lowest orbit of the old quantum theory) with frequency ω. The same formula, (1·0), was obtained, with the difference that $\gamma = 2{\cdot}42$.

These classical calculations, however, were entirely upset by the old quantum theory. For in the case of an atom at some distance from the track of the α-particle, the transfer of energy is less than the energy of ionisation, and must have certain discrete values, namely the differences between the energies of the various quantum orbits. Henderson‖ suggested that when the classical transfer lies between two of the discrete values, the actual transfer must be taken to be the smaller of these. The result was, naturally, a smaller stopping power than before, not more than half the experimental value¶.

* Thomson, *Conduction of Electricity through Gases*, pp. 370–382.
† Darwin, *Phil. Mag.* XXIII, p. 901 (1912).
‡ Bohr, *Phil. Mag.* XXV, p. 10 (1913).
§ Fowler, *Proc. Camb. Phil. Soc.* XXII, p. 793 (1925).
‖ Henderson, *Phil. Mag.* XLIV, p. 680 (1922).
¶ See Fowler, *Proc. Camb. Phil. Soc.* XXI, p. 521 (1923).

Digitized by Google

The present paper contains a rough application of the new quantum theory to the less close collisions. It does not treat the α-particle as part of the dynamical system, nor replace it by de Broglie waves; but represents it by a point moving with constant velocity in a straight line, and forming the centre of a Coulomb field. The distance of the track from the atomic nucleus (supposed fixed) is large compared with atomic dimensions. The effect on the atom can then be found by perturbation theory, which must be carried to the second order, since the first order gives a zero result.

The new quantum theory deals in statistical averages, and gives a mean energy transfer without any hypothesis about individual jumps. The surplus energy over the energy of a jump is therefore not wasted as in Henderson's theory. On taking over from the previous theories the transfer by close collisions, which has not yet been computed on the new theory, we arrive again at the formula (1·0), with the following modifications:

γ has Bohr's value 1·123,

$\omega = 1{\cdot}10\omega'$, where ω' is the frequency of the head of the Lyman series.

Thus we may write

$$-\frac{dT}{dx} = \frac{4\pi N \mathrm{E}^2\epsilon^2}{\mu v^2} \log \frac{\gamma' \mu v^3}{2\pi\omega' \mathrm{E}\epsilon} \quad \ldots\ldots\ldots\ldots(1{\cdot}1),$$

where $\gamma' = 1{\cdot}02$.

This is the formula for monatomic hydrogen, for which, of course, no experimental data exist. For comparison with observation, it is necessary to extend the formulae (1·0), (1·1) to the case of hydrogen molecules and helium. In Bohr's theory the formula for such gases is

$$-\frac{dT}{dx} = \frac{8\pi N \mathrm{E}^2\epsilon^2}{\mu v^2} \log \frac{\gamma \mu v^3}{2\pi \mathrm{E}\epsilon\omega} \ldots\ldots\ldots\ldots(1{\cdot}2).$$

Fowler* takes over this formula, and assumes that the ω's for various molecules are proportional to their ionisation potentials. He shows that the best empirical value for γ is about 1. If we make a similar modification of (1·1),

$$-\frac{dT}{dx} = \frac{8\pi N \mathrm{E}^2\epsilon^2}{\mu v^2} \log \frac{\gamma' \mu v^3}{2\pi\omega' \mathrm{E}\epsilon} \quad \ldots\ldots\ldots\ldots(1{\cdot}3),$$

ω' is now the frequency corresponding to the ionisation potential of H_2 or He, and is half the ω of (1·2). The optimum value of γ' is therefore $\frac{1}{2}$, and not 1·02, as in (1·1). If we retain the value $\gamma' = 1{\cdot}02$, the stopping power of H_2 at N.T.P. for $v = 1{\cdot}98 \cdot 10^9$ cm. sec.$^{-1}$ is $6{\cdot}16 \cdot 10^{35}$. Gurney's† experimental value is $5{\cdot}4 \cdot 10^{35}$.

* Fowler, *Proc. Camb. Phil. Soc.* XXII, p. 794 (1925).

† Gurney, *Proc. Roy. Soc.* A, CVII, pp. 332, 340 (1925).

Digitized by Google

The ratio of the stopping powers of H_2 and He at this velocity is 1·17, instead of 1·19.

However, the step from (1·1) to (1·3) is a long and uncertain one. It will be seen that the present theory gives γ' in a very complicated manner, and there seems to be no reason to believe that it will have the same value for H_2 and He as for H, though we may be fairly confident of the form of (1·3).

The point which this paper is intended to bring out is that on the new quantum theory, as on the classical theory, there is an average transfer of energy from the α-particle to atoms at all distances, and that of nearly the classical amount; so that the difficulties in the theory of the stopping of α-particles are no longer fundamental, as in the old quantum theory, but are mere complications of algebra.

2. *General Theory**.

We have to deal with a special case of the following general problem: a number of equivalent dynamical systems (atoms) have initially a known energy-distribution; what will be the energy-distribution after a small variable perturbing potential has acted upon each for a given time? In our special case, the dynamical systems are H-atoms, initially all in their normal state; the perturbing field is that of the α-particle; and we require the proportion of the atoms with any given energy after the encounter.

We start with Schrödinger's equation in the form

$$\left[H\left(q, \frac{h}{\iota}\frac{\partial}{\partial q}\right) - \frac{h}{\iota}\frac{\partial}{\partial t}\right]\psi = 0 \quad \text{.....(2·00),}$$

where the unperturbed Hamiltonian is supposed independent of the time and h is Planck's constant divided by 2π.

For the unperturbed system we may put

$$\psi = ue^{\frac{\iota W}{h}t} \quad \text{.....(2·01),}$$

where u is independent of t, and satisfies

$$\left[H\left(q, \frac{h}{\iota}\frac{\partial}{\partial q}\right) - W\right]u = 0 \quad \text{.....(2·02).}$$

The condition that u is single-valued and continuous introduces quantum conditions in the well-known manner. In general, W will have a set of discrete characteristic values, and also a continuous range of values, which satisfy the condition. We shall suppose that the discrete characteristic values are expressed in

* All this section is parallel to Dirac, *Proc. Roy. Soc.* A, CXII, pp. 673–677. See also Born, *Zeitschrift für Physik*, 40, p. 167.

Digitized by Google

terms of quantum numbers $(n, k, m \ldots)$, which also specify the corresponding functions u; and that in the continuous range the functions u are specified by W and by quantum numbers $(k, m \ldots)$.

The functions $u(n, k, m \ldots)$ are normalised in the ordinary sense of the word by the choice of suitable constant multipliers. That is

$$\int u\,(n, k, m \ldots)\, u^*(n', k', m' \ldots)\, d\tau = \begin{cases} 1 \text{ if } n = n',\ k = k',\ m = m' \ \ldots \\ 0 \text{ otherwise} \end{cases} \qquad \ldots\ldots(2{\cdot}1),$$

where $d\tau$ is a generalised volume element of the configuration-space, over the whole of which the integral is taken, and u^* is the conjugate complex number to u.

The functions $u(W, k, m \ldots)$ must also be multiplied by a suitable factor. But they are not normalised in the sense of (2·1). The factor is chosen so that an arbitrary function f of the q's can be expanded in the form†

$$f = \sum_{n, k, m \ldots} \gamma(n, k, m \ldots)\, u(n, k, m \ldots) + \int \sum_{k, m \ldots} \gamma(W, k, m \ldots)\, u(W, k, m \ldots)\, \frac{dW}{2\pi h} \qquad \ldots\ldots(2{\cdot}20),$$

where the $\int$ is taken over the continuous range of characteristics,

$$\gamma(n, k, m \ldots) = \int f u^*(n, k, m \ldots)\, d\tau \qquad \ldots\ldots(2{\cdot}21),$$

and‡

$$\gamma(W, k, m \ldots) = \lim_{\Delta \to 0} \frac{1}{\Delta} \int f \int_{W}^{W + 2\pi h\Delta} u^*(W, k, m \ldots)\, \frac{dW}{2\pi h}\, d\tau \ldots(2{\cdot}22)$$

$$= \int f u^*(W, k, m \ldots)\, d\tau \qquad \ldots\ldots\ldots\ldots\ldots\ldots(2{\cdot}23),$$

if this integral is uniformly convergent.

We shall replace such equations as (2·20) by

$$f = (\Sigma_s + \int)\, \gamma(s)\, u(s) \qquad \ldots\ldots\ldots\ldots\ldots(2{\cdot}24),$$

where s summarises the parameters $(n, k, m \ldots)$ or $(W, k, m \ldots)$.

We shall suppose such conditions of convergence in our sums and integrals that (2·23) is valid, and that the orders of sums and integrals and differentiations can be inverted.

Initially, our set of dynamical systems has the wave-function

$$\psi_0 = (\Sigma_s + \int)\, c(s)\, \psi(s) \qquad \ldots\ldots\ldots\ldots\ldots(2{\cdot}30),$$

where $c(s)$ is constant, and $\psi(s)$ is given by (2·01).

For the discrete characteristics, Dirac has shown that

$$P(s) = c(s)\, c^*(s) \qquad \ldots\ldots\ldots\ldots\ldots(2{\cdot}31)$$

† See Oppenheimer, *Proc. Camb. Phil. Soc.* XXIII, p. 422 (1926).
‡ See Fues, *Ann. der Physik*, LXXXI, 19, pp. 295–303 (1926).

Digitized by Google

is the proportion of the dynamical systems which are in the state denoted by s. It is natural to extend (2·31) to the continuous range of W, with the following interpretation of P:

The proportion of systems with energies between W and $W + dW$, and with quantum numbers $(k, m \ldots)$, is

$$P(W, k, m \ldots)\, dW/2\pi h \quad \ldots\ldots\ldots\ldots\ldots\ldots (2{\cdot}32).$$

This interpretation is in harmony with Oppenheimer's paper; also it will be seen later that $(\Sigma_s + \int) P(s)$ remains constant (namely unity) in spite of perturbations.

We now introduce a perturbing potential $A(t)$ which is a function of the time as well as of the coordinates, and is supposed to be small. Schrödinger's equation becomes

$$\left[H\left(q, \frac{h}{\iota}\frac{\partial}{\partial q}\right) + A(t) - \frac{h}{\iota}\frac{\partial}{\partial t}\right]\psi = 0 \quad \ldots\ldots\ldots (2{\cdot}40).$$

By (2·24), at any given time T, ψ can be expressed in the form

$$\psi = (\Sigma_s + \int)\, \gamma(s, T)\, u(s),$$

where γ is constant for given s, T.

Evidently we may write in general

$$\psi = (\Sigma_s + \int)\, a(s, t)\, \psi(s) \quad \ldots\ldots\ldots\ldots\ldots (2{\cdot}41),$$

where $a(s, t)$ depends on the time and not on the coordinates.

The variable proportion of systems in a given state is then given by

$$P(s, t) = a(s, t)\, a^*(s, t) \quad \ldots\ldots\ldots\ldots\ldots (2{\cdot}42)$$

(see (2·31)).

The P's, which give the energy-distribution of the dynamical systems, are therefore determined by the a's.

Now substitute from (2·41) in (2·40), remembering that $\psi(s)$ satisfies (2·00),

$$(\Sigma_s + \int) A(t)\, a(s, t)\, \psi(s) - (\Sigma_s + \int)\frac{h}{\iota}\dot{a}(s, t)\, \psi(s) = 0 \quad \ldots (2{\cdot}50).$$

Let $$A(sr;\ t) \equiv \int A(t)\, u(s)\, u^*(r)\, d\tau \ldots\ldots\ldots\ldots (2{\cdot}51),$$

the integral being taken over the whole configuration-space.

Then by (2·21), (2·23) and (2·24):

$$A(t)\, u(s) = (\Sigma_r + \int) A(sr;\ t)\, u(r) \quad \ldots\ldots\ldots (2{\cdot}52).$$

Substitute in (2·50)

$$(\Sigma_s + \int)(\Sigma_r + \int)\, a(s, t)\, A(sr;\ t)\, u(r)\, e^{\iota \frac{W(s)}{h} t}$$
$$- (\Sigma_s + \int)\frac{h}{\iota}\dot{a}(s, t)\, u(s)\, e^{\iota \frac{W(s)}{h} t} = 0.$$

Digitized by Google

Interchanging the pair of $(\Sigma + \int)$'s, we have a series of the form (2·24) equated to 0. Its coefficients may be presumed to vanish:

$$(\Sigma_r + \int)\, a(r,t)\, A(rs;\ t)\, e^{\iota \frac{W(r)}{h} t} - \frac{h}{\iota}\, \dot{a}(s,t)\, e^{\iota \frac{W(s)}{h} t} = 0,$$

i.e. $$\dot{a}(s,t) = \frac{\iota}{h}(\Sigma_r + \int)\, a(r,t)\, A(rs;\ t)\, e^{\iota \frac{W(r) - W(s)}{h} t} \quad \ldots(2{\cdot}53).$$

Write, for brevity,

$$A'(rs;\ t) = A(rs;\ t)\, e^{\iota \frac{W(r) - W(s)}{h} t} \quad \ldots\ldots\ldots(2{\cdot}54).$$

Then $$\dot{a}(s,t) = \frac{\iota}{h}(\Sigma_r + \int)\, a(r,t)\, A'(rs;\ t) \quad \ldots\ldots\ldots(2{\cdot}55).$$

This equation determines the a's, since their initial values are known:

at $t = t_0$ $$a(s, t_0) = c(s) \quad \ldots\ldots\ldots\ldots\ldots\ldots(2{\cdot}56).$$

The right-hand side of (2·55) is proportional to the perturbing potential, and is therefore small. We solve the equation by successive approximations, going as far as the second.

Let $$a(s,t) = c(s) + a_1(s,t) + a_2(s,t),$$

where a_1, a_2 are small quantities of the first and second orders respectively. The first approximation is given by

$$\left.\begin{aligned} \dot{a}_1(s,t) &= \frac{\iota}{h}(\Sigma_r + \int)\, c(r)\, A'(rs;\ t) \\ a_1(s, t_0) &= 0 \end{aligned}\right\},$$

therefore $$a_1(s,t) = \frac{\iota}{h}(\Sigma_r + \int)\, c(r) \int_{t_0}^{t} A'(rs;\ \theta)\, d\theta \ldots(2{\cdot}60).$$

For the second approximation,

$$\left.\begin{aligned} \dot{a}_2(s,t) &= \frac{\iota}{h}(\Sigma_r + \int)\, \frac{\iota}{h}(\Sigma_u + \int)\, c(u) \int_{t_0}^{t} A'(ur;\ \theta)\, d\theta \,.\, A'(rs;\ t) \\ a_2(s, t_0) &= 0 \end{aligned}\right\},$$

therefore

$$a_2(s,t_1) = -\frac{1}{h^2}(\Sigma_u + \int)\, c(u)\, (\Sigma_r + \int) \int_{t_0}^{t_1} dt \int_{t_0}^{t} d\theta\, A'(ur;\ \theta)\, A'(rs;\ t) \quad \ldots\ldots(2{\cdot}61).$$

Write, for brevity,

$$\left.\begin{aligned} \alpha(rs) &= \frac{\iota}{h} \int_{t_0}^{t_1} A'(rs;\ \theta)\, d\theta \\ \beta(us) &= -\frac{1}{h^2}(\Sigma_r + \int) \int_{t_0}^{t_1} dt \int_{t_0}^{t} d\theta\, A'(ur;\ \theta)\, A'(rs;\ t) \end{aligned}\right\} \ldots(2{\cdot}62),$$

so that α, β are of the first and second orders respectively.

Digitized by Google

Then

$$a(s, t_1) = c(s) + a_1(s, t_1) + a_2(s, t_1)$$
$$= c(s) + (\Sigma_r + \int) c(r)\, \alpha(rs) + (\Sigma_u + \int) c(u)\, \beta(us) \ldots (2{\cdot}63).$$

Substitute in (2·42), neglecting small quantities of order higher than the second,

$$P(s, t_1) = a(s, t_1)\, a^*(s, t_1)$$
$$= c(s)\, c^*(s) + (\Sigma_r + \int)\, [c(s)\, c^*(r)\, \alpha^*(rs) + c^*(s)\, c(r)\, \alpha(rs)]$$
$$+ (\Sigma_r + \int)(\Sigma_u + \int)\, c(r)\, c^*(u)\, \alpha(rs)\, \alpha^*(us)$$
$$+ (\Sigma_u + \int)\, [c(s)\, c^*(u)\, \beta^*(us) + c^*(s)\, c(u)\, \beta(us)]$$
$$\ldots\ldots (2{\cdot}70).$$

This expresses $P(s, t_1)$ in terms of the c's; but the c's are not completely determined by the initial values of the P's. As can be seen from (2·31), the argument of $c(s)$ is quite unknown. We assume therefore that all arguments (phases) occur with equal frequency, and average over all their values. This can be done by substituting $c(s)\, e^{i\gamma_s}$ for $c(s)$ and integrating with respect to the γ's. The result is to leave only terms of the form $c(s)\, c^*(s)$, which can be replaced by the initial P's. After averaging,

$$P(s, t_1) = P(s, t_0) + P(s, t_0)\, [\alpha^*(ss) + \alpha(ss)]$$
$$+ (\Sigma_r + \int)\, P(r, t_0)\, \alpha(rs)\, \alpha^*(rs) + P(s, t_0)\, [\beta^*(ss) + \beta(ss)]$$
$$\ldots\ldots (2{\cdot}71).$$

Now, by (2·51) and (2·54),

$$A'^*(rs;\ t) = A'(sr;\ t),$$

therefore, by (2·62), $\alpha^*(rs) = -\alpha(sr)$,

therefore
$$\left.\begin{aligned} \alpha^*(ss) + \alpha(ss) &= 0 \\ \alpha(rs)\, \alpha^*(rs) &= -\alpha(rs)\, \alpha(sr) \end{aligned}\right\} \ldots\ldots\ldots\ldots (2{\cdot}72).$$

Again,

$$\beta^*(ss) + \beta(ss)$$
$$= -\frac{1}{h^2}(\Sigma_r + \int) \left[\int_{t_0}^{t_1} dt \int_{t_0}^{t} d\theta\, A'(rs, \theta)\, A'(sr, t) \right.$$
$$\left. + \int_{t_0}^{t_1} dt \int_{t_0}^{t} d\theta\, A'(sr, \theta)\, A'(rs, t) \right]$$
$$= -\frac{1}{h^2}(\Sigma_r + \int) \int_{t_0}^{t_1} A'(rs, t)\, dt \int_{t_0}^{t_1} A'(sr, t)\, dt$$
$$= (\Sigma_r + \int)\, \alpha(rs)\, \alpha(sr) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2{\cdot}73).$$

Writing $\Delta P(s)$ for $P(s, t_1) - P(s, t_0)$,
$P(s)$ for $P(s, t_0)$,

(2·71) becomes

$$\Delta P(s) = \frac{1}{h^2}(\Sigma_r + \int)\, [P(r) - P(s)] \left| \int_{t_0}^{t_1} A'(sr;\ t)\, dt \right|^2 \ldots (2{\cdot}74).$$

Digitized by Google

This gives the required answer. Note that a first approximation alone gives a zero result, owing to the first equation of (2·72). Also that

$$(\Sigma_s + \int) \Delta P(s) \equiv 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots (2{\cdot}8),$$

showing that our physical interpretation of P is plausible.

3. *The Reaction between an α-Particle and a Hydrogen Atom.*

The nucleus of the atom is supposed fixed at the origin of coordinates.

$-\epsilon$, μ are the charge and mass of the electron.

$a_0 \equiv h^2/\mu\epsilon^2$, the radius of the innermost orbit in Bohr's theory.

The α-particle is supposed to move with a constant velocity v in a straight track, whose distance p from the nucleus is large compared with a_0; and to give rise to an electric potential E/R, where E is the charge on the α-particle, and R the distance from its instantaneous position.

The atom is initially in its normal state.

We use spherical polar coordinates (r, θ, ϕ), the axis, $\theta = 0$, being perpendicular to the plane containing the nucleus and the track, and the plane $\phi = 0$ perpendicular to the track. The origin of time is the instant at which the α-particle crosses the plane $\phi = 0$.

For the unperturbed hydrogen atom, Schrödinger's* equation (2·02) becomes

$$\left[-\frac{h^2}{2\mu}\nabla^2 - \frac{\epsilon^2}{r} - W\right] u = 0 \quad \ldots\ldots\ldots\ldots (3{\cdot}00).$$

The characteristic values of W form a discrete set of negative values, and the whole range of positive numbers. The discrete values are given by

$$W(n) = -\frac{\mu\epsilon^4}{2h^2n^2} = -\frac{\epsilon^2}{2a_0n^2} \quad \ldots\ldots\ldots\ldots (3{\cdot}01),$$

where n must be a positive integer.

In the continuous range, we shall have occasion to use a number n' given by

$$W = +\frac{\mu\epsilon^4}{2h^2n'^2} = +\frac{\epsilon^2}{2a_0n'^2} \quad \ldots\ldots\ldots\ldots (3{\cdot}02),$$

where n' is positive, but not necessarily integral.

The normalised solutions of (3·00) for the negative values of W (elliptic orbits) are†

$$u(nkm) = C(nkm)\, e^{im\Phi} P_k^m(\cos\theta)\, \alpha^k e^{-\alpha/2} L_{n+k}^{2k+1}(\alpha) \ldots (3{\cdot}03),$$

* Schrödinger, *Ann. der Physik*, LXXIX, p. 361 (1926).

† See Schrödinger, *loc. cit.*

Digitized by Google

where k is a positive integer less than n, or zero,

m is an integer numerically $\leqslant k$,

$$P_k^m(\cos\theta) = \sin^m\theta \frac{d^m}{(d\cos\theta)^m} P_k(\cos\theta)$$

($P_k(x)$ is a Legendre polynomial),

$$L_{n+k}^{2k+1}(\alpha) = \frac{d^{2k+1}}{d\alpha^{2k+1}} L_{n+k}(\alpha)$$

($L_{n+k}(\alpha)$ is a Laguerre polynomial),

$$\alpha = \frac{2r}{na_0},$$

$C(nkm)$ is a normalising factor

.....(3·04).

The solutions of (3·00) for positive W, normalised in the sense of (2·20), are*

$$u(Wkm) = C(Wkm)\, e^{im\phi} P_k^m(\cos\theta)\, r^k \int_{(-\gamma,\gamma)} e^{zr} (z-\gamma)^{\alpha_1 - 1} (z+\gamma)^{\alpha_2 - 1}\, dz$$

.....(3·05),

where k is any positive integer, or zero,

m is an integer numerically $\leqslant k$,

the complex integral is taken along a double loop enclosing $-\gamma$, γ,

$$\alpha_1 = k+1-\iota n', \qquad \alpha_2 = k+1+\iota n', \qquad \gamma = \iota/n'a_0,$$

$$C(Wkm) = \left[\frac{2k+1}{4\pi}\frac{(k-m)!}{(k+m)!}\right]^{\frac{1}{2}} \times \frac{\sqrt{\frac{2\mu}{h}}\, |2\gamma|^{-k-\frac{1}{2}} e^{\frac{3\pi}{2}n'}}{(1-e^{2\pi\iota\alpha_1})(1-e^{2\pi\iota\alpha_2})\sqrt{\{\Gamma(\alpha_1)\,\Gamma(\alpha_2)\}}}$$

.....(3·06).

The perturbing potential is

$$A(t) = -\frac{E\epsilon}{R} \qquad \text{.....(3·10).}$$

We shall denote the normal state of the atom by a '1.' The initial proportions (at $t = -\infty$) are

$$\left.\begin{aligned} P(1) &= 1 \\ P(s) &= 0 \text{ for other states} \end{aligned}\right\} \qquad \text{.....(3·20).}$$

Hence, if in (2·74) s denotes a state other than the normal state,

$$\Delta P(s) = \frac{1}{h^2}\left|\int_{-\infty}^{\infty} A'(s1;t)\, dt\right|^2 \qquad \text{.....(3·21).}$$

* See Oppenheimer, *loc. cit.* The second half of $C(Wkm)$ was communicated to me by Dr Oppenheimer, by letter.

Digitized by Google

$\Delta P(1)$ can be found, if required, from (2·8), which shows that the loss to the normal state is equal to the gain to the other states.

We have now to evaluate $A'(s1; t)$ separately for the cases in which the state s has a negative energy (excitation), and for those in which it has a positive energy (ionisation).

4. *Excitation.*

The normal state has quantum numbers (1, 0, 0).
Let the state s have quantum numbers (n, k, m).
By (2·54) and (2·51), and (3·10),

$$A'(s1; t) \equiv A'(nkm; 100)$$

$$= e^{i\frac{W(n)-W(1)}{h}t} \int -\frac{\mathrm{E}\epsilon}{R} u(nkm)\, u^*(100)\, d\tau \ldots(4{\cdot}00),$$

where $d\tau \equiv r^2 \sin\theta\, dr\, d\theta\, d\phi$, and the integral is taken through all space.

Substituting from (3·03),

$$A'(nkm; 100) = \mathrm{E}\epsilon\, C(nkm)\, C(100)\, e^{i\frac{W(n)-W(1)}{h}t}$$

$$\times \int \frac{e^{im\phi} P_k^m(\cos\theta)\, e^{-\frac{n+1}{2}\alpha} \alpha^k L_{n+k}^{2k+1}(\alpha)\, d\tau}{R} \ldots(4{\cdot}01),$$

where $$\alpha = \frac{2r}{na_0};$$

for $$P_0^0(\cos\theta) = 1, \quad L_1^1\left(\frac{2r}{a_0}\right) = -1.$$

Now the integral in (4·01) is the electrostatic or gravitational potential at the position of the α-particle of a volume density represented by the numerator; that is, the value at this point of a potential V which vanishes at infinity, is everywhere twice differentiable, and satisfies the equation

$$\nabla^2 V = -4\pi e^{im\phi} P_k^m(\cos\theta)\, e^{-\frac{n+1}{2}\alpha} \alpha^k L_{n+k}^{2k+1}(\alpha) \ldots(4{\cdot}02).$$

Such a potential V can be found without difficulty.

Put $$V = -4\pi e^{im\phi} P_k^m(\cos\theta) f(\alpha) \left(\frac{na_0}{2}\right)^2 \quad \ldots\ldots(4{\cdot}03).$$

Then

$$\left[\frac{d^2}{d\alpha^2} + \frac{2}{\alpha}\frac{d}{d\alpha} - \frac{k(k+1)}{\alpha^2}\right] f(\alpha) = e^{-\frac{n+1}{2}\alpha} \alpha^k L_{n+k}^{2k+1}(\alpha)$$

$$\ldots\ldots(4{\cdot}04).$$

Digitized by Google

The equation (4·04) has the following solution:

$$f(\alpha) = \frac{-1}{2k+1}\left[\alpha^k \int_a^\infty e^{-\frac{n+1}{2}x} x L_{n+k}^{2k+1}(x)\,dx + \alpha^{-(k+1)} \int_0^a e^{-\frac{n+1}{2}x} x^{2(k+1)} L_{n+k}^{2k+1}(x)\,dx\right] \quad \ldots(4{\cdot}05),$$

which behaves like α^k at the origin, and like $\alpha^{-(k+1)}$ at infinity.

The integral in (4·01) is thus

$$-4\pi\left(\frac{na_0}{2}\right)^2 \left(\frac{p+\iota vt}{\sqrt{(p^2+v^2t^2)}}\right)^m P_k^m(0) f\left(\frac{2\sqrt{(p^2+v^2t^2)}}{na_0}\right) \quad \ldots\ldots(4{\cdot}06).$$

Write $$z = \frac{vt}{p}, \quad u = \frac{W(n)-W(1)}{h}\frac{p}{v} \quad \ldots\ldots\ldots\ldots(4{\cdot}10).$$

Then, by (4·01), (4·06),

$$\int_{-\infty}^{\infty} A'(nkm;\,100)\,dt = -\mathrm{E}\epsilon C(nkm)\,C(100)\,4\pi\left(\frac{na_0}{2}\right)^2 P_k^m(0) \times \int_{-\infty}^{\infty} e^{\iota uz}\left(\frac{1+\iota z}{\sqrt{(1+z^2)}}\right)^m f\left(\frac{2p\sqrt{(1+z^2)}}{na_0}\right)\frac{p}{v}\,dz \ldots(4{\cdot}11).$$

If $m \geqslant 0$, this may be written

$$\left.\begin{aligned}\int_{-\infty}^{\infty} A'(n,k,m;\,100)\,dt &= -\mathrm{E}\epsilon C(nkm)\,C(100)\,4\pi\left(\frac{na_0}{2}\right)^2 P_k^m(0)\frac{p}{v}\left(1+\frac{\partial}{\partial u}\right)^m \\ &\quad \times \int_{-\infty}^{\infty} \frac{e^{\iota uz} f\left(\frac{2p\sqrt{(1+z^2)}}{na_0}\right) dz}{(1+z^2)^{m/2}}\end{aligned}\right\}$$

Similarly

$$\left.\begin{aligned}\int_{-\infty}^{\infty} A'(n,k,-m;\,100)\,dt &= -\mathrm{E}\epsilon C(nkm)\,C(100)\,4\pi\left(\frac{na_0}{2}\right)^2 P_k^m(0)\frac{p}{v}\left(1-\frac{\partial}{\partial u}\right)^m \\ &\quad \times \int_{-\infty}^{\infty} \frac{e^{\iota uz} f\left(\frac{2p\sqrt{(1+z^2)}}{na_0}\right) dz}{(1+z^2)^{m/2}}\end{aligned}\right\} \quad \ldots\ldots(4{\cdot}12).$$

Digitized by Google

It is now necessary to consider the relative magnitudes of the various terms in the integral in (4·12).

$f(\alpha)$ contains some terms of the form

$$\text{constant } \alpha^q e^{-\frac{n+1}{2}\alpha} \qquad (-k-1 \leqslant q \leqslant n) \ldots\ldots (4{\cdot}20),$$

and a single term containing no exponent, which may be shewn to be

$$\frac{kn}{2k+1}\frac{(n+k)!^2}{(n-k-1)!}\frac{2^{2k+4}(n-1)^{n-k-2}}{(n+1)^{n+k+2}}\alpha^{-k-1} \ldots\ldots (4{\cdot}21).$$

The term in the integral arising from (4·20) is proportional to

$$\lambda^q \int_{-\infty}^{\infty} (1+z^2)^{\frac{q-m}{2}} e^{\iota u z - \lambda\sqrt{(1+z^2)}}\, dz \ldots\ldots (4{\cdot}22),$$

where $$\lambda = \frac{(n+1)p}{na_0}.$$

Since λ is large, the integrand of (4·22) is negligible except for small z*; thus (4·22) may be replaced by

$$e^{-\lambda}\lambda^q \int_{-\infty}^{\infty} e^{\iota u z - \frac{\lambda}{2}z^2} dz$$

$$= e^{-\lambda}\lambda^q \int_{-\infty+\frac{\iota u}{\lambda}}^{\infty+\frac{\iota u}{\lambda}} e^{\iota u z - \frac{\lambda}{2}z^2} dz$$

(since the integrand is regular, and vanishes rapidly at ∞ between the two paths of integration)

$$= e^{-\left(\lambda+\frac{u^2}{2\lambda}\right)}\lambda^q \int_{-\infty}^{\infty} e^{-\frac{\lambda}{2}x^2} dx$$

$$= \sqrt{2\pi}\lambda^{q-\frac{1}{2}} e^{-\left(\lambda+\frac{u^2}{2\lambda}\right)} \ldots\ldots (4{\cdot}23).$$

The term in the integral arising from (4·21) is

$$\frac{k}{2k+1}\frac{(n+k)!^2}{(n-k-1)!}\frac{2^{k+3}n^{k+2}(n-1)^{n-k-2}}{(n+1)^{n+k+2}}\left(\frac{a_0}{p}\right)^{k+1}\int_{-\infty}^{\infty}\frac{e^{\iota u z}\, dz}{(1+z^2)^{\frac{k+m+1}{2}}}$$

$\ldots\ldots (4{\cdot}24).$

The last integral is expressible in terms of a Bessel function†,

$$\int_{-\infty}^{\infty}\frac{e^{\iota u z}\, dz}{(1+z^2)^{\frac{k+m+1}{2}}} = 2\frac{\pi^{\frac{1}{2}} u^{\frac{k+m}{2}} K_{k+m}(u)}{2^{\frac{k+m}{2}}\Gamma\left(\frac{k+m+1}{2}\right)\cos\frac{k+m}{2}\pi} \qquad (4{\cdot}25).$$

* This approximation may break down when u is nearly as large as λ. If so, we must use the fact that u/λ is, in practice, small—see the text following equation (4·28).

† Watson, *Theory of Bessel Functions*, Camb. Univ. Press, 1922, p. 185.

Digitized by Google

Also*, as $u \to \infty$,

$$K_{\frac{k+m}{2}}(u) \sim \left(\frac{\pi}{2u}\right)^{\frac{1}{2}} e^{-u} \cos\frac{k+m}{2}\pi \quad \ldots\ldots\ldots(4{\cdot}26).$$

It follows that as $p \to \infty$, (4·24) behaves like

$$\frac{e^{-\frac{W(n)-W(1)}{h\nu}p}}{p^{\frac{k-m+3}{2}}},$$

and since m cannot be greater than k, the states for which $m = k$ have a much stronger excitation than the rest. For such states (4·24) becomes

$$\frac{k}{2k+1}\frac{(n+k)!^2}{(n-k-1)!}\frac{2^{k+4}n^{k+2}(n-1)^{n-k-2}}{(n+1)^{n+k+2}} \times \frac{a_0^{k+1}}{p}\left(\frac{W(n)-W(1)}{h\nu}\right)^k \frac{K_k(u)}{(-)^k(2k-1)!} \quad \ldots\ldots(4{\cdot}27).$$

Having regard to (4·26), (4·27) is larger than (4·23), since

$$\lambda + \frac{u^2}{2\lambda} = u + \tfrac{1}{2}\lambda + \frac{(u-\lambda)^2}{2\lambda} > u + \tfrac{1}{2}\lambda.$$

Thus (4·23) may be neglected.

What remains of (4·12) is

$$\int_{-\infty}^{\infty} A'(n,k,\pm k;100)\,dt = -\mathrm{E}\epsilon C(nkk)C(100)\,4\pi\left(\frac{na_0}{2}\right)^2(2k-1)!\frac{p}{v} \times \left(1 \pm \frac{\partial}{\partial u}\right)^k \frac{u^k K_k(u)}{(-)^k(2k-1)!}\left(\frac{a_0}{p}\right)^{k+1}\frac{k}{2k+1}\frac{(n+k)!^2}{(n-k-1)!} \times \frac{2^{k+4}n^{k+2}(n-1)^{n-k-2}}{(n+1)^{n+k+2}} \quad \ldots\ldots\ldots(4{\cdot}28).$$

The expression (4·28) contains the term

$$\left(\frac{ua_0}{p}\right)^k = \left(\frac{W(n)-W(1)}{h\nu}a_0\right)^k.$$

Now for reasonable values of v, nearing 2×10^9 cms. sec.$^{-1}$,

$$\frac{W(n)-W(1)}{h\nu}a_0 < \frac{-W(1)}{h\nu}a_0 = \text{about } 0{\cdot}055.$$

The rest of the factors in (4·28) vary little from one value to another of k, except that the expression vanishes when $k = 0$. Thus its value for $k = 1$ is roughly 20 times that for $k = 2$, and

* Watson, p. 202.

Digitized by Google

so on. Since (4·28) is to be squared for insertion in (3·21), we need retain only the states for which $k = 1$.

$$\int_{-\infty}^{\infty} A'(n, 1, \pm 1; 100)\, dt = -\,\mathrm{E}\epsilon C(n11)\, C(100)\, 4\pi \left(\frac{na_0}{2}\right)^2 \frac{p}{v} \times \left(1 \pm \frac{\partial}{\partial u}\right) \frac{uK_1(u)}{-1} \left(\frac{a_0}{p}\right)^2 \cdot \frac{1}{3} \frac{(n+1)!^2}{(n-2)!} \frac{2^5 n^3 (n-1)^{n-3}}{(n+1)^{n+3}} \quad \ldots\ldots(4{\cdot}29).$$

In this expression

$$\left(1 \pm \frac{\partial}{\partial u}\right) uK_1(u) = u[K_1(u) \mp K_0(u)] \quad \ldots\ldots(4{\cdot}30).$$

By (3·03) and (3·04)

$$\frac{1}{C^2(n11)} = \int_0^{2\pi} d\phi \int_0^{\pi} [P_1^1(\cos\theta)]^2 \sin\theta\, d\theta \int_0^{\infty} \alpha^2 e^{-\alpha} [L^3_{n+1}(\alpha)]^2 r^2\, dr$$

$$= 2\pi \cdot \frac{4}{3} \left(\frac{na_0}{2}\right)^3 2n^2 (n^2-1)(n+1)!^2 * \quad \ldots\ldots(4{\cdot}31).$$

Therefore

$$C(n11) = \frac{1}{(n+1)!\sqrt{\{\frac{2}{3}\pi a_0^3 n^5 (n^2-1)\}}},$$

$$\frac{1}{C^2(100)} = \int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta\, d\theta \int_0^{\infty} e^{-2r/a_0} r^2\, dr$$

$$= 2\pi \cdot 2 \cdot \left(\frac{a_0}{2}\right)^3 \cdot 2 \quad \ldots\ldots(4{\cdot}32),$$

$$C(100) = \frac{1}{\sqrt{(\pi a_0^3)}}.$$

Substituting in (4·29),

$$\int_{-\infty}^{\infty} A'(n, 1, \pm 1; 100)\, dt = \frac{2^4\sqrt{2}}{\sqrt{3}} \frac{\mathrm{E}\epsilon a_0}{pv} \frac{n^{\frac{7}{2}} (n-1)^{n-\frac{5}{2}}}{(n+1)^{n+\frac{1}{2}}} u[K_1(u) \mp K_0(u)] \quad \ldots\ldots(4{\cdot}33).$$

Substituting in (3·21),

$$\Delta P(n, 1, \pm 1) = \frac{2^9}{3} \left(\frac{\mathrm{E}\epsilon a_0}{pvh}\right)^2 \frac{n^7 (n-1)^{2n-5}}{(n+1)^{2n+3}} u^2 [K_1(u) \mp K_0(u)]^2.$$

Owing to the degeneracy of the hydrogen atom these quantities separately have no physical importance. Their sum, however, is the proportion of atoms which acquire an energy corresponding to the total quantum number n.

$$\Delta P(n) = \frac{2^{10}}{3} \left(\frac{\mathrm{E}\epsilon a_0}{pvh}\right) \frac{n^7 (n-1)^{2n-5}}{(n+1)^{2n+3}} u^2 [K_1^2(u) + K_0^2(u)] \quad \ldots\ldots(4{\cdot}4).$$

* See Schrödinger, *Zeit. f. Phys.* LXXX, p. 437 (1926), Mathematischer Anhang.

Digitized by Google

5. *Ionisation.*

Let the state s have positive energy W, and quantum numbers k, m.

Corresponding to (4·00), we have

$$A'(s1;\ t) = A'(Wkm;\ 100)$$

$$= e^{\iota \frac{W - W(1)}{h} t} \int - \frac{\mathrm{E}\epsilon}{R}\, u(Wkm)\, u^*(100)\, d\tau \quad \ldots(5{\cdot}00),$$

where $d\tau = r^2 \sin\theta\, dr\, d\theta\, d\phi$, and the integral is taken through all space.

Substituting from (3·05),

$$A'(Wkm;\ 100) = \mathrm{E}\epsilon\, C(Wkm)\, C(100)\, e^{\iota \frac{W - W(1)}{h} t}$$

$$\times \int \frac{e^{\iota m\phi} P_k^m(\cos\theta)\, r^k e^{-r/a} \int_{(-\gamma,\,\gamma)} e^{zr} (z-\gamma)^{a_1 - 1} (z+\gamma)^{a_2 - 1}\, dz}{R}\, d\tau$$

$$\ldots\ldots(5{\cdot}01).$$

If we choose the path of the complex integral so that on it

$$Rz < \tfrac{1}{2} a_0 \quad \ldots\ldots\ldots\ldots\ldots\ldots(5{\cdot}02),$$

we may invert the two integrals; for the space-integral so obtained is uniformly convergent for the path of the complex integral.

Thus

$$A'(Wkm;\ 100) = \mathrm{E}\epsilon\, C(Wkm)\, C(100)\, e^{\iota \frac{W - W(1)}{h} t}$$

$$\times \int_{(-\gamma,\,\gamma)} \int \frac{e^{\iota m\phi} P_k^m(\cos\theta)\, r^k e^{-r(1/a_0 - z)}\, d\tau}{R} (z-\gamma)^{a_1 - 1} (z+\gamma)^{a_2 - 1}\, dz$$

$$\ldots\ldots(5{\cdot}03).$$

The space-integral in (5·03) is evaluated in the same way as in the last section. It is the value at the position of the α-particle of a potential function V, which vanishes at infinity, is twice differentiable everywhere, and satisfies

$$\nabla^2 V = -4\pi e^{\iota m\phi} P_k^m(\cos\theta)\, r^k e^{-r(1/a_0 - z)} \quad \ldots\ldots(5{\cdot}04).$$

The solution of (5·04) is

$$V = -4\pi e^{\iota m\phi} P_k^m(\cos\theta)\, g(r, z) \quad \ldots\ldots\ldots(5{\cdot}05),$$

where

$$g(r, z) = \frac{-1}{2k+1} \left[r^k \int_r^\infty e^{-(1/a_0 - z)x}\, x\, dx + r^{-(k+1)} \int_0^r e^{-(1/a_0 - z)x}\, x^{2(k+1)}\, dx \right]$$

$$\ldots\ldots(5{\cdot}06).$$

Digitized by Google

Thus the integral in (5·03) is

$$\int_{(-\gamma,\gamma)} -4\pi\left(\frac{p+\iota vt}{\sqrt{(p^2+v^2t^2)}}\right)^m P_k^m(0)\, g(\sqrt{(p^2+v^2t^2)}, z) \times (z-\gamma)^{a_1-1}(z+\gamma)^{a_2-1}\, dz \quad \ldots(5{\cdot}07).$$

Write $$\theta = \frac{vt}{p},\quad u = \frac{W-W(1)}{h}\frac{p}{v} \quad \ldots\ldots\ldots\ldots\ldots(5{\cdot}10).$$

Then, by (5·03), (5·07),

$$\int_{-\infty}^{\infty} A'(Wkm;\ 100)\, dt = -4\pi \mathrm{E}\epsilon\, C(Wkm)\, C(100)\, P_k^m(0) \times \int_{-\infty}^{\infty} e^{\iota u\theta}\left(\frac{1+\iota\theta}{\sqrt{(1+\theta^2)}}\right)^m \int_{(-\gamma,\gamma)} g(p\sqrt{(1+\theta^2)}, z) \times (z-\gamma)^{a_1-1}(z+\gamma)^{a_2-1}\, dz\, \frac{p}{v}\, d\theta \quad \ldots\ldots\ldots\ldots\ldots(5{\cdot}11).$$

With the proviso (5·02), the order of integration may be inverted, and if $m \geqslant 0$ we may write

$$\int_{-\infty}^{\infty} A'(W, k, \pm m;\ 100)\, dt = -4\pi \mathrm{E}\epsilon\, C(Wkm)\, C(100)\, P_k^m(0)\, \frac{p}{v} \times \left(1 \pm \frac{\partial}{\partial u}\right)^m \int_{(-\gamma,\gamma)} \int_{-\infty}^{\infty} \frac{e^{\iota u\theta}\, g(p\sqrt{(1+\theta^2)}, z)}{(1+\theta^2)^{m/2}}\, d\theta \times (z-\gamma)^{a_1-1}(z+\gamma)^{a_2-1}\, dz \quad \ldots\ldots\ldots\ldots\ldots(5{\cdot}12).$$

As before, we consider the magnitude of the various terms arising from g.

With the proviso (5·02), the terms containing exponential factors may again be neglected. The remaining term in (5·06) is

$$-\frac{1}{2k+1}\frac{(2k+2)!}{(1, a_0-z)^{k+2}}\, r^{-k-1}.$$

Its contribution to the inner integral in (5·12) is

$$-\frac{1}{2k+1}\frac{(2k+2)!}{(1, a_0-z)^{k+2}}\frac{1}{p^{k-1}}\int_{-\infty}^{\infty} \frac{e^{\iota u\theta}\, d\theta}{(1+\theta^2)^{\frac{k+m+1}{2}}} = -\frac{1}{2k+1}\frac{(2k+2)!}{(1, a_0-z)^{k+2}}\frac{u^{\frac{k+m}{2}}}{p^{k+1}}\frac{2\pi^{\frac{1}{2}} K_{\frac{k+m}{2}}(u)}{2^{\frac{k+m}{2}}\, \Gamma\left(\frac{k+m+1}{2}\right)\cos\frac{k+m}{2}\pi} \quad \ldots\ldots(5{\cdot}20).$$

Digitized by Google

As before, the states for which $k = m$ are more prominent than the others. By substitution of (5·20) in (5·12),

$$\int_{-\infty}^{\infty} A'(W, k, \pm k;\ 100)\, dt = 4\pi \mathrm{E}\epsilon\, C(Wkk)\, C(100)\, \frac{p}{v}$$
$$\times \frac{2.(2k+2)!}{(-)^k (2k+1)} \left(1 \pm \frac{\partial}{\partial u}\right)^k \frac{u^k}{p^{k+1}} K_k(u)$$
$$\times \int_{(-\gamma,\ \gamma)} (1/a_0 - z)^{-(2k+3)} (z-\gamma)^{a_1-1} (z+\gamma)^{a_2-1}\, dz \qquad \text{.....(5·21).}$$

By taking W large enough, $\frac{ua_0}{p}$ may be made as large as we like; so that our previous argument for retaining only the smallest possible values of k breaks down for large energies. In these cases, however, u being large, $K_k(u)$ is very small, by (4·26); so that such energies may be left out of account entirely.

For $k = 0$, the complex integral vanishes. The important terms are those for which $k = 1$,

$$\int_{-\infty}^{\infty} A'(W, 1, \pm 1;\ 100)$$
$$= -2^6 \pi \frac{\mathrm{E}\epsilon}{pv}\, C(W11)\, C(100)\, u\,(K_1(u) \mp K_0(u))$$
$$\times \int_{(-\gamma,\ \gamma)} (1/a_0 - z)^{-5} (z-\gamma)^{a_1-1} (z+\gamma)^{a_2-1}\, dz \text{.....(5·22),}$$

using (4·30).

We now evaluate the complex integral. By (3·06), for $k = 1$,

$$a_1 = 2 - \iota n', \quad a_2 = 2 + \iota n', \quad \gamma = \iota/n' a_0.$$

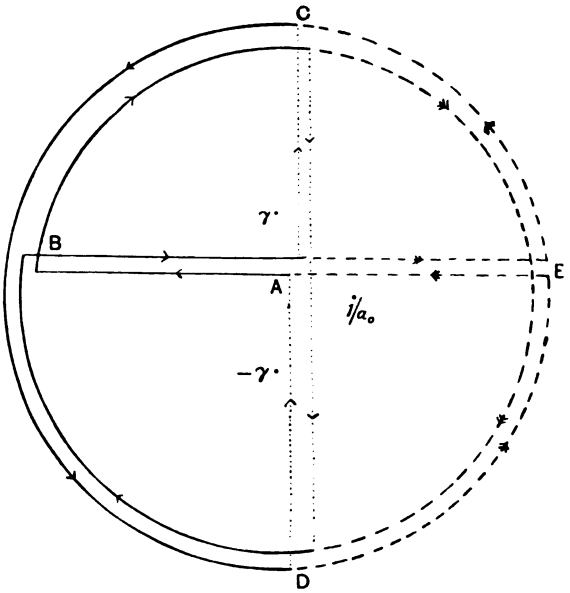
The integral is

$$\int_{\left(-\frac{\iota}{na_0},\ \frac{\iota}{na_0}\right)} \left(\frac{1}{a_0} - z\right)^{-5} \left(z - \frac{\iota}{na_0}\right)^{1-\iota n'} \left(z + \frac{\iota}{na_0}\right)^{1+\iota n'} dz$$
$$\text{.....(5·30).}$$

Having regard to (5·02), a suitable path of integration is that drawn in the figure, partly in full line and partly dotted, namely $ABCDBACBDA$. The curved parts ideally coincide on a circle whose radius tends to infinity. The horizontal lines coincide, and pass between the real axis and $+\gamma$; the vertical lines coincide and pass between the imaginary axis and $1/2a_0$.

Digitized by Google

We compare the integral (5·30) with a second integral, whose integrand is the same, but whose path of integration consists of the full and dashed parts of the figure, namely $ABCEDBAECBDEA$.



The contribution to this second integral of the circular parts of the path of integration vanishes as the radius of the circle tends to infinity. The integrals along BAE and EAB cancel each other. For in passing from one of these parts of the path to the other, each of the points $\gamma, -\gamma$ is encircled once; so that the integrand is altered in the ratio

$$e^{2\pi\iota(1-\iota n')}e^{2\pi\iota(1+\iota n')} = e^{4\pi\iota} = 1.$$

That is, the integrand is the same in both cases, while the line is described in opposite senses.

The integral, then, over the full and dashed path vanishes; so that (5·30) is equal to its excess over that integral. This is equivalent to the integral of the same expression over the three dotted and dashed circuits $ACEA$, $AEDA$, $ADECA$.

Inside the first circuit the integrand is regular; inside the other two there is a singularity at $1/a_0$. The second circuit is clockwise, the third counter-clockwise. Also in passing from a point on $AEDA$ to a point on $ADECA$, along the original path of

Digitized by Google

integration, γ is encircled once clockwise, and $-\gamma$ not at all; so that the ratio of the integrands is

$$1 : e^{-2\pi\iota(1-\iota n')} = e^{2\pi n'} : 1.$$

Hence the integral (5·30) is equal to

$$(1 - e^{2\pi n'})\, 2\pi\iota \times \text{residue at } 1/a_0 \quad \ldots\ldots\ldots\ldots (5{\cdot}31)$$

for that branch of the integrand which corresponds to CAD.

The residue is found to be

$$-\tfrac{2}{3}\frac{n'^4 a_0^2}{(n'^2+1)^2} e^{-2n'(\pi+\cot^{-1}n')} \quad \ldots\ldots\ldots\ldots (5{\cdot}32),$$

where $0 \leqslant \cot^{-1} n' \leqslant \pi/2$.

From (3·06)

$$C(W11) = \frac{\sqrt{\dfrac{3\mu}{4\pi h}}\left(\dfrac{2}{n' a_0}\right)^{-\frac{5}{2}} e^{\frac{3\pi}{2}n'}}{(1 - e^{2\pi n'})(1 - e^{-2\pi n'})\sqrt{(\Gamma(2-\iota n')\,\Gamma(2+\iota n'))}} \quad \ldots\ldots (5{\cdot}33),$$

and

$$\begin{aligned}\Gamma(2-\iota n')\,\Gamma(2+\iota n') &= (1-\iota n')\,\Gamma(1-\iota n').\,(1+\iota n')\,\iota n'\,\Gamma(\iota n') \\ &= \iota n'(n'^2+1)\frac{\pi}{\sin \iota n'\pi} = \frac{2\pi n'(n'^2+1)}{e^{n'\pi} - e^{-n'\pi}},\end{aligned}$$

therefore

$$C(W11) = -\frac{1}{8\pi}\sqrt{\frac{3\mu a_0^2}{h}}\frac{n'}{\sqrt{(n'^2+1)}}\frac{e^{\frac{3\pi}{2}n'}}{(e^{n'\pi} - e^{-n'\pi})^{\frac{3}{2}}} \quad \ldots (5{\cdot}34).$$

From (4·32)

$$C(100) = \frac{1}{\sqrt{(\pi a_0^3)}} \quad \ldots\ldots (5{\cdot}35).$$

Substituting (5·31), (5·32), (5·34), (5·35) in (5·22),

$$\int_{-\infty}^{\infty} A'(W, 1, \pm 1; 100) = \frac{2^5\sqrt{\pi}}{\sqrt{3}}\iota\frac{\mathrm{E}\epsilon a_0^2}{pv}\sqrt{\frac{\mu}{h}}\frac{n'^5}{(n'^2+1)^{\frac{5}{2}}}\frac{e^{-2n'\cot^{-1}n'}}{\sqrt{(1-e^{-2\pi n'})}} \times u[K_1(u) \mp K_0(u)] \ldots\ldots (5{\cdot}40).$$

Substituting in (3·21),

$$\begin{aligned}&\Delta P(W, 1, \pm 1) \\ &= \frac{2^{10}\pi}{3}\left(\frac{\mathrm{E}\epsilon a_0}{pvh}\right)^2 \frac{\mu a_0^2}{h}\frac{n'^{10}}{(n'^2+1)^5}\frac{e^{-4n'\cot^{-1}n'}}{1-e^{-2\pi n'}} u^2[K_1(u) \mp K_0(u)]^2 \\ &\qquad \ldots\ldots (5{\cdot}41),\end{aligned}$$

and

$$\Delta P(W) = \frac{2^{11}\pi}{3}\left(\frac{\mathrm{E}\epsilon a_0}{pvh}\right)^2 \frac{\mu a_0^2}{h}\frac{n'^{10}}{(n'^2+1)^5}\frac{e^{-4n'\cot^{-1}n'}}{1-e^{-2\pi n'}} u^2[K_1^2(u) + K_0^2(u)] \quad \ldots\ldots (5{\cdot}42).$$

Digitized by Google

It is worth noticing that, on the border-line between excitation and ionisation, the equations (4·4) and (5·42) are equivalent. For as $W \to -0$, $n \to \infty$ by (3·01); as $W \to +0$, $n' \to \infty$ by (3·02).

Now (4·4) gives, for the proportion of atoms in states with energies between W and $W + dW$, in the limit when n is large,

$$\Delta P(n) \left| \frac{dn}{dW} \right| dW$$

$$= \frac{2^{10}}{3} \left(\frac{\mathrm{E}\epsilon a_0}{pvh}\right)^2 \frac{n^7 (n-1)^{2n-5}}{(n+1)^{2n+5}} u^2 [K_1^2(u) + K_0^2(u)] \frac{h^2 n^3}{\mu\epsilon^4} dW$$

(by (3·01))

$$\to \frac{2^{10}}{3} \left(\frac{\mathrm{E}\epsilon a_0}{pvh}\right)^2 \frac{h^2}{\mu\epsilon^4} e^{-4u^2} [K_1^2(u) + K_0^2(u)] dW \quad \ldots(5{\cdot}50),$$

while (5·52) gives, by (2·32),

$$\Delta P(W) \frac{dW}{2\pi h}$$

$$\to \frac{2^{10}}{3} \left(\frac{\mathrm{E}\epsilon a_0}{pvh}\right)^2 \frac{\mu a_0^2}{h^2} e^{-4u^2} [K_1^2(u) + K_0^2(u)] dW \ldots(5{\cdot}51).$$

The value of u in each of these two expressions is (by (4·10) and (5·10))

$$u = -\frac{W(1)}{h} \frac{p}{v},$$

and by the definition of a_0 (beginning of § 3),

$$\frac{\mu a_0^2}{h^2} = \frac{h^2}{\mu\epsilon^4},$$

so that (5·50) and (5·51) are equal for equal ranges of energy.

6. *The Calculation of the Stopping Power.*

The mean energy lost by an α-particle in passing a hydrogen atom at a distance p large compared with atomic dimensions is

$$-\Delta T = (\Sigma_s + \int) \Delta P(s) [W(s) - W(1)] \ldots\ldots(6{\cdot}00),$$

where ΔP is given approximately by (4·4) and (5·42).

The question arises whether this expression can be fitted on to the formula for close collisions (small p),

$$\left.\begin{aligned} -\Delta T &= \frac{2\mathrm{E}^2\epsilon^2}{\mu v^2 (p^2 + \lambda^2)} \\ \text{where} \quad \lambda &= \frac{\mathrm{E}\epsilon}{\mu v^2} \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots(6{\cdot}01).$$

Digitized by Google

This formula was calculated by Thomson* on the assumption that for close collisions the electron could be treated as free. It has figured in every subsequent theory, and can probably be taken over into this one.

Now as $p \to 0$, or $v \to \infty$, $u \to 0$, and $u^2[K_1^2(u) + K_0^2(u)] \to 1$. Thus the limiting value of (6·00) is

$$-\Delta T \to \sum_{n=2}^{\infty} \left(\frac{\mathrm{E}\epsilon a_0}{pvh}\right)^2 \frac{2^{10}}{3} \frac{n^7(n-1)^{2n-5}}{(n+1)^{2n+5}} [W(n) - W(1)]$$

$$+ \int_0^{\infty} \left(\frac{\mathrm{E}\epsilon a_0}{pvh}\right)^2 \frac{\mu a_0^2}{h} \frac{2^{11}\pi}{3} \frac{n'^{10}}{(n'^2+1)^5} \frac{e^{-4n'\cot^{-1}n'}}{1-e^{-2\pi n'}} [W - W(1)] \frac{dW}{2\pi h}$$

$$= \frac{2\mathrm{E}^2\epsilon^2}{\mu v^2 p^2} \left[\frac{2^8}{3} \sum_{n=2}^{\infty} \frac{n^5(n-1)^{2n-4}}{(n+1)^{2n+4}} + \frac{2^8}{3} \int_0^{\infty} \frac{n'^5}{(n'^2+1)^4} \frac{e^{-4n'\cot^{-1}n'}}{1-e^{-2\pi n'}} dn'\right]$$

.....(6·02),

using the definition of a_0, (3·01) and (3·02).

I have calculated numerically the bracket in (6·02), and obtained the result 0·992. I have little doubt that the exact value is unity. If that be so, (6·00) and (6·01) agree when $v \to \infty$; and (6·00) for small values of p merges into (6·01) for (comparatively) large values of p. Following Bohr, we shall assume that the whole range can be covered with a good approximation by

$$-\Delta T = \frac{p^2}{p^2+\lambda^2} (\Sigma_s + \int) \Delta P(s) [W(s) - W(1)] \quad ...(6·03).$$

This reduces to (6·01) for small p, and to (6·00) for large p.

Now if there are N atoms per unit volume in the gas, the number which the α-particle will pass in unit length of its track at a distance between p and $p + dp$ is

$$N \cdot 2\pi p\, dp.$$

So the rate of loss of energy per unit length of track is

$$-\frac{dT}{dx} = \int_0^{\infty} \frac{p^2}{p^2+\lambda^2} (\Sigma_s + \int) \Delta P(s) [W(s) - W(1)] N \cdot 2\pi p\, dp$$

.....(6·10).

We can integrate under the $(\Sigma + \int)$. The terms in $\Delta P(s)$ containing p are, by (4·4) and (5·42),

$$\frac{u^2}{p^2} [K_1^2(u) + K_0^2(u)],$$

where, by (4·10) and (5·10),

$$u = \frac{W - W(1)}{h} \frac{p}{v}.$$

* Thomson, *loc. cit.*

Digitized by Google

Thus we require

$$\int_0^\infty \frac{u^2}{p^2+\lambda^2}[K_1^2(u)+K_0^2(u)]\,p\,dp$$

$$=\int_0^\infty \frac{u^2}{u^2+\lambda'^2}[K_1^2(u)+K_0^2(u)]\,u\,du \quad \text{.....(6·11),}$$

where $\lambda' = \dfrac{W-W(1)}{h\nu}$ $\lambda = \dfrac{E\epsilon}{\mu h \nu^3}(W-W(1))$.

λ' is small, so that (6·11) may be split into two integrals, which are approximately

$$\int_0^\eta \frac{u\,du}{u^2+\lambda'^2} + \int_\eta^\infty [K_1^2(u)+K_0^2(u)]\,u\,du \quad \text{.....(6·12),}$$

where η is small, but large compared with λ'.

Also $$[K_1^2(u)+K_0^2(u)]\,u = -\frac{d}{du}[uK_1(u)\,K_0(u)],$$

therefore (6·12) is

$$\log\eta - \log\lambda' - \Big[uK_1(u)\,K_0(u)\Big]_\eta^\infty \quad \text{.....(6·13).}$$

Now, since η is small*,

$$\eta K_1(\eta)\sim 1, \quad K_0(\eta)\sim -(\log\tfrac{1}{2}\eta + 0{\cdot}5772),$$

therefore (6·13) is approximately

$$-0{\cdot}5772 + \log 2 - \log\lambda' = \log\gamma - \log\frac{E\epsilon}{\mu\nu^3} - \log\frac{W-W(1)}{h}$$

.....(6·14),

where $\gamma = 1{\cdot}123$ nearly.

Thus the integral (6·11) is nearly

$$\log\frac{\gamma\mu\nu^3 h}{-E\epsilon\,W(1)} - \log\frac{W-W(1)}{-W(1)} \quad \text{.....(6·15).}$$

Substituting in (6·10), and reducing in the same way as for (6·02),

$$-\frac{dT}{dx} = \frac{4\pi N E^2\epsilon^2}{\mu\nu^2}\left\{\log\frac{\gamma\mu\nu^3 h}{-E\epsilon\,W(1)}\left[\frac{2^3}{3}\sum_{n=2}^{\infty}\frac{n^5(n-1)^{2n-4}}{(n+1)^{2n+4}}\right.\right.$$

$$+\frac{2^3}{3}\int_0^\infty \frac{n'^5}{(n'^2+1)^4}\,\frac{e^{-4n'\cot^{-1}n'}}{1-e^{-2\pi n'}}\,dn'\Big] - \left[\frac{2^3}{3}\sum_{n=2}^{\infty}\frac{n^5(n-1)^{2n-4}}{(n+1)^{2n+4}}\log\frac{n^2-1}{n^2}\right.$$

$$\left.\left.+\frac{2^3}{3}\int_0^\infty \frac{n'^5}{(n'^2+1)^4}\,\frac{e^{-4n'\cot^{-1}n'}}{1-e^{-2\pi n'}}\log\frac{n'^2+1}{n'^2}\,dn'\right]\right\} \quad \text{.....(6·20).}$$

* Watson, p. 80.

Digitized by Google

The second square bracket in (6·20) was also evaluated numerically, giving $0{\cdot}10_0 = \log 1{\cdot}10_5$ nearly. Thus

$$-\frac{dT}{dx} = \frac{4\pi N \mathrm{E}^2\epsilon^2}{\mu v^2} \log \frac{\gamma \mu v^3}{2\pi \mathrm{E}\epsilon\omega} \quad \text{.....(6·21),}$$

where

$$\omega = -1{\cdot}10\,\frac{W(1)}{2\pi h} \quad \text{.....(6·22).}$$

Since $2\pi h$ is Planck's constant, $-\dfrac{W(1)}{2\pi h}$ is the frequency of the head of the spectral series associated with the normal state; or in other words, the frequency corresponding to the ionisation potential. If we call this ω', and put

$$\gamma' = \frac{\gamma}{1{\cdot}10} \quad \text{.....(6·23),}$$

(6·21) becomes

$$-\frac{dT}{dx} = \frac{4\pi N \mathrm{E}^2\epsilon^2}{\mu v^2} \log \frac{\gamma' \mu v^3}{2\pi\omega' \mathrm{E}\epsilon}.$$

This is the formula (1·1) discussed in the introduction.

I should like to thank Mr R. H. Fowler for suggesting this problem, and for his indispensable advice at all stages.

Digitized by Google

Digitized by Google

Digitized by Google

PROCEEDINGS
OF THE
Cambridge Philosophical Society.

A case of distinction between Fourier integrals and Fourier series.

By MARGARET ELEANOR GRIMSHAW, Newnham College.
(Communicated by Mr S. POLLARD.)

[*Received* 12 March, *read* 2 May, 1927.]

1. A Fourier integral is said to be of *finite type** if its generating function vanishes for all sufficiently large values of $|x|$. Because the coefficient functions are defined by integrals over a finite range, the behaviour of such a Fourier integral usually resembles closely that of the corresponding series.

It may be shewn that when $k \geqslant 1$ the necessary and sufficient conditions for the summability (C, k) of a Denjoy-Fourier integral of finite type are exactly the same as those for the corresponding series. When $k < 1$ this is no longer the case, and the integral may be summable while the series is not. Indeed, it is possible to give an example of a function whose series is summable (C, k) for $0 \leqslant k \leqslant \mathbf{k} < 1$ almost nowhere in the interval $(-\pi, \pi)$, while the integral is summable for all $k > 0$ everywhere in the interval.

This result is due to the fact that for the coefficients of a Denjoy-Fourier series the most that can generally be stated is

$$a_n = o(n), \quad b_n = o(n),$$

so that the series is, in general, at most summable $(C, 1)$, whereas the corresponding integral is not necessarily so restricted by the behaviour of the coefficients.

To make matters quite definite, let $f(x)$ be zero outside $(-\pi, \pi)$ and integrable in the general Denjoy sense in $(-\pi, \pi)$. Then $f(x)$ has a Fourier series

$$\tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx), \tag{1·1}$$

where

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos nt\,dt, \quad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin nt\,dt \quad (n = 0, 1, 2, \ldots), \tag{1·2}$$

and it has a Fourier integral of finite type

$$\int_0^{\infty} (a_s \cos sx + b_s \sin sx)\,ds, \tag{1·3}$$

where

$$a_s = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos st\,dt, \quad b_s = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin st\,dt \quad (s \geqslant 0). \tag{1·4}$$

* S. Pollard, *Proc. Camb. Phil. Soc.*, 23 (1926), 373—382.

Digitized by Google

For a particular value of $k \geqslant 1$ the summability (C, k) at a point x in $-\pi < x < \pi$ of either the integral or the series implies that of the other, and the sums are the same. For other values of k the conditions differ in quite a marked degree.

2. In a consideration of the summability of the integral (1·3) we are led to investigate the behaviour of

$$(2\cdot1) \qquad I_k(\omega) = \int_0^\omega \left(1 - \frac{s}{\omega}\right)^k (a_s \cos sx + b_s \sin sx)\, ds$$

as $\omega \to \infty$.

The corresponding expression for the series (1·1) is

$$(2\cdot2) \qquad \sigma_k(\omega) = \tfrac{1}{2}a_0 + \sum_{1 \leqslant n < \omega} \left(1 - \frac{n}{\omega}\right)^k (a_n \cos nx + b_n \sin nx).$$

The final form of the criteria for summability of the integral depends essentially upon the possibility of changing the order of integration in a repeated integral, and is in some respects easier to obtain than the corresponding form for the series.

It will be noticed that $I_k(\omega)$ does not exist, in general, unless $k > -1$. In what follows this will be assumed throughout, and other restrictions on k will be introduced as required.

For the purpose of comparison we consider a generating function $f(x)$ which vanishes outside the interval $(-\pi, \pi)$. It will be evident, however, that the results of paragraphs 3 and 5 are quite general, and can be established in exactly the same way for a function which is integrable in the general Denjoy sense in any finite interval (p, q) and vanishes outside this interval.

3. THEOREM I. *If* $k > -1$, *then*

$$(3\cdot1) \qquad I_k(\omega) = \frac{\omega}{\pi} \int_{-\pi-x}^{\pi-x} f(x+t)\, \gamma(\omega t)\, dt,$$

where

$$(3\cdot2) \qquad \gamma(t) = \int_0^1 (1-u)^k \cos tu\, du.$$

It immediately appears that

$$\omega\gamma(\omega t) = \int_0^\omega \left(1 - \frac{u}{\omega}\right)^k \cos tu\, du$$

so that (3·1) is equivalent to

$$I_k(\omega) = \frac{1}{\pi} \int_{-\pi-x}^{\pi-x} dt f(x+t) \int_0^\omega \left(1 - \frac{s}{\omega}\right)^k \cos ts\, ds.$$

Digitized by Google

Moreover, by (1·4) and (2·1),

$$I_k(\omega) = \frac{1}{\pi}\int_0^\omega ds\left(1-\frac{s}{\omega}\right)^k \int_{-\pi}^{\pi} f(t)\cos s(t-x)\,dt$$

$$= \frac{1}{\pi}\int_0^\omega ds\left(1-\frac{s}{\omega}\right)^k \int_{-\pi-x}^{\pi-x} f(x+t)\cos st\,dt. \tag{3·3}$$

Thus the result to be established is equivalent to the possibility of inverting the order of integration in (3·3).

Write $$g(x) = f(x) - \tfrac{1}{2}a_0,$$

so that, by (1·2), $$\int_{-\pi}^{\pi} g(t)\,dt = 0.$$

If $$G(x) = \int_{-\pi}^{x} g(t)\,dt,$$

then $$G(-\pi) = G(\pi) = 0.$$

By integration by parts, which is justified since $\cos st$ is of bounded variation in any finite interval,

$$\int_{-\pi-x}^{\pi-x} g(x+t)\cos st\,dt = \Big[G(x+t)\cos st\Big]_{-\pi-x}^{\pi-x} + s\int_{-\pi-x}^{\pi-x} G(x+t)\sin st\,dt$$

$$= s\int_{-\pi-x}^{\pi-x} G(x+t)\sin st\,dt.$$

Thus (3·3) becomes

$$I_k(\omega) = \frac{a_0}{2\pi}\int_0^\omega ds\left(1-\frac{s}{\omega}\right)^k \int_{-\pi-x}^{\pi-x}\cos st\,dt + \frac{1}{\pi}\int_0^\omega ds\,s\left(1-\frac{s}{\omega}\right)^k \int_{-\pi-x}^{\pi-x} G(x+t)\sin st\,dt.$$

Now $G(x+t)$ is continuous, and so $|G(x+t)|$ has a finite upper bound M in $(-\pi-x, \pi-x)$. If we replace the integrands in the above integrals by their absolute values, the result is not greater than

$$a_0\int_0^\omega \left(1-\frac{s}{\omega}\right)^k ds + 2\omega M\int_0^\omega \left(1-\frac{s}{\omega}\right)^k ds < \infty$$

since $k > -1$.

Digitized by Google

By de la Vallée Poussin's theorem we may change the order of integration and write

$$I_k(\omega) = \frac{a_0}{2\pi}\int_{-\pi-x}^{\pi-x} dt \int_0^\omega \left(1-\frac{s}{\omega}\right)^k \cos ts\,ds$$

$$+ \frac{1}{\pi}\int_{-\pi-x}^{\pi-x} dt\, G(x+t)\int_0^\omega s\left(1-\frac{s}{\omega}\right)^k \sin ts\,ds.$$

Now

$$\int_{-\pi-x}^{\pi-x} dt\, g(x+t)\int_0^\omega \left(1-\frac{s}{\omega}\right)^k \cos ts\,ds$$

$$= \left[G(x+t)\int_0^\omega \left(1-\frac{s}{\omega}\right)^k \cos ts\,ds\right]_{-\pi-x}^{\pi-x}$$

$$+ \int_{-\pi-x}^{\pi-x} dt\, G(x+t)\int_0^\omega s\left(1-\frac{s}{\omega}\right)^k \sin ts\,ds$$

$$= \int_{-\pi-x}^{\pi-x} dt\, G(x+t)\int_0^\omega s\left(1-\frac{s}{\omega}\right)^k \sin ts\,ds,$$

the integration by parts being allowed since

$$\int_0^\omega \left(1-\frac{s}{\omega}\right)^k \cos ts\,ds$$

regarded as a function of t is of bounded variation in any finite interval.

Thus

$$I_k(\omega) = \frac{1}{\pi}\int_{-\pi-x}^{\pi-x} dt\,\{\tfrac{1}{2}a_0 + g(x+t)\}\int_0^\omega \left(1-\frac{s}{\omega}\right)^k \cos ts\,ds$$

$$= \frac{\omega}{\pi}\int_{-\pi-x}^{\pi-x} f(x+t)\,\gamma(\omega t)\,dt,$$

which establishes the theorem.

4. For $k > -1$, $\qquad \int_0^1 (1-u)^k\,du$

is absolutely convergent.

Also $$\cos tu = \sum_{n=0}^{\infty} (-)^n \frac{(ut)^{2n}}{(2n)!},$$

the series being, for each t, uniformly convergent in u in $(0, 1)$.

Digitized by Google

Thus by (3·2)

$$\gamma(t) = \sum_{n=0}^{\infty} (-)^n \frac{t^{2n}}{(2n)!} \int_0^1 (1-u)^k u^{2n}\, du$$

$$= \sum_{n=0}^{\infty} (-)^n \frac{t^{2n}}{(2n)!} \frac{\Gamma(k+1)\,\Gamma(2n+1)}{\Gamma(2n+k+2)}$$

$$= \frac{\Gamma(k+1)}{t^{k+1}} \sum_{n=0}^{\infty} (-)^n \frac{t^{2n+k+1}}{\Gamma(2n+k+2)}$$

$$= \frac{\Gamma(k+1)}{t^{k+1}} C_{k+1}(t),$$

where $$C_p(t) = \sum_{n=0}^{\infty} (-)^n \frac{t^{p+2n}}{\Gamma(p+2n+1)}.$$

This last function has been investigated in detail* and many of its properties are well known.

For $t \geqslant 0$ and $0 \leqslant p \leqslant 2$, $C_p(t)$ is bounded, while for $p > 2$

$$C_p(t) \sim \frac{t^{p-2}}{\Gamma(p+1)}$$

as $t \to \infty$.

It follows that

$$|\gamma(t)| \leqslant \frac{A}{|t|^{k+1}} \qquad (-1 < k \leqslant 1) \tag{4·1}$$

$$\gamma(t) \sim \frac{k}{t^2} \qquad (k \geqslant 1), \tag{4·2}$$

where here, and in subsequent contexts, A is constant relative to the function considered.

Now $C_p(t)$ is defined as above for $p > -1$, and it is clear that

$$C_p(t) = \frac{t^p}{\Gamma(p+1)} - C_{p+2}(t).$$

Thus $C_p(t)$ is bounded for $t \geqslant 1$ and $-1 < p \leqslant 2$.

Moreover $$\frac{d}{dt} C_{p+1}(t) = C_p(t),$$

and so $$\frac{1}{\Gamma(k+1)} \gamma'(t) = -\frac{k+1}{t} \frac{C_{k+1}(t)}{t^{k+1}} + \frac{1}{t} \frac{C_k(t)}{t^k};$$

whence, for $t \geqslant 1$,

$$|\gamma'(t)| \leqslant \frac{A}{|t|^{1+\alpha}} \qquad (k > -1) \tag{4·3}$$

where $\alpha = \min(k, 2)$.

* W. H. Young, *Quart. Journal of Maths.*, 43 (1912), 161—177.

Digitized by Google

For small values of t, $\gamma(t)$ and all its differential coefficients are bounded, since

$$\gamma^{(r)}(t) = \int_0^1 (1-u)^k u^r \cos^{(r)} tu\, du \quad (r = 0, 1, 2, \ldots),$$

and therefore

$$\gamma^{(r)}(t) \leqslant \int_0^1 (1-u)^k u^r\, du = B(k+1, r+1).$$

This may be written

$$|\gamma^{(r)}(t)| \leqslant A. \tag{4·4}$$

From (4·3) and (4·4) it immediately appears that

$$\int_0^\infty |\gamma'(t)|\, dt < \infty$$

provided that $k > 0$. That is, $\gamma(t)$ is of bounded variation in $(0, \infty)$ if $k > 0$.

It is also evident from (4·4) that $\gamma(t)$ is of bounded variation in any finite range for all $k > -1$.

5. When $k > -1$ the limits of $I_k(\omega)$ are identical with those of

$$\frac{\omega}{\pi} \int_{-\pi-x}^{\pi-x} f(x+t)\, \gamma(\omega t)\, dt.$$

If now η is any positive number less than both $\pi - x$ and $\pi + x$,

$$\int_\eta^{\pi-x} f(x+t)\, \omega\gamma(\omega t)\, dt = \Big[F(x+t)\, \omega\gamma(\omega t)\Big]_\eta^{\pi-x} - \int_\eta^{\pi-x} F(x+t)\, \omega^2 \gamma'(\omega t)\, dt, \tag{5·1}$$

where $F(t)$ is an indefinite integral of $f(t)$, the integration by parts being justified since $\gamma(t)$ is of bounded variation in any finite range.

In the range of integration

$$|\omega\gamma(\omega t)| \leqslant \frac{A}{\omega^k \eta^{k+1}} \quad (-1 < k \leqslant 1)$$

$$\leqslant \frac{A}{\omega\eta^2} \qquad (k \geqslant 1)$$

by (4·1) and (4·2).

Take now $k > 0$, and consider the right-hand side of (5·1). The term between limits tends to zero with $\omega\gamma(\omega t)$ as $\omega \to \infty$.

Digitized by Google

For the integral, if (λ, μ) be any interval in $(\eta, \pi - x)$,

$$\int_\lambda^\mu \omega^2 \gamma'(\omega t)\, dt = \Big[\omega\gamma(\omega t)\Big]_\lambda^\mu \to 0,$$

so that, since $F(x+t)$ is continuous, by the Riemann-Lebesgue convergence theorem

$$\int_\eta^{\pi-x} F(x+t)\, \omega^2 \gamma'(\omega t)\, dt \to 0$$

as $\omega \to \infty$.

Thus
$$\int_\eta^{\pi-x} f(x+t)\, \omega\gamma(\omega t)\, dt \to 0.$$

In like manner

$$\int_{-\pi-x}^{-\eta} f(x+t)\, \omega\gamma(\omega t)\, dt \to 0,$$

and the limits of $I_k(\omega)$, for $k > 0$, are the same as those of

$$\frac{\omega}{\pi}\int_{-\eta}^{\eta} f(x+t)\, \omega\gamma(\omega t)\, dt.$$

It is known that

(5·2)
$$\frac{1}{\pi}\int_0^\infty \gamma(t)\, dt = \tfrac{1}{2}\,*.$$

Thus, as $\omega \to \infty$,
$$\frac{\omega}{\pi}\int_0^\eta \gamma(\omega t)\, dt \to \tfrac{1}{2}.$$

If then we write

$$\phi(t) = f(x+t) + f(x-t) - 2s$$

the necessary and sufficient condition for summability (C, k) of the integral (1·3) to the sum s at x takes the form

$$\frac{\omega}{\pi}\int_0^\eta \phi(t)\, \gamma(\omega t)\, dt \to 0.$$

By what appears above the limits of this last integral are independent of η, and the condition may be written

(5·3)
$$\lim_{\eta\to 0}\ \overline{\lim_{\omega\to\infty}}\ \left|\omega\int_0^\eta \phi(t)\, \gamma(\omega t)\, dt\right| = 0 \quad (k > 0).$$

Write now
$$\Phi(t) = \int_0^t \phi(u)\, du.$$

* W. H. Young, *Quart. Journal of Maths.*, 43 (1912), 166, 170.

Digitized by Google

Integration by parts gives

$$\omega \int_0^\eta \phi(t)\,\gamma(\omega t)\,dt = \Big[\omega\Phi(t)\,\gamma(\omega t)\Big]_0^\eta - \omega^2 \int_0^\eta \Phi(t)\,\gamma'(\omega t)\,dt$$

$$= \omega\Phi(\eta)\,\gamma(\omega\eta) - \omega^2 \int_0^\eta \Phi(t)\,\gamma'(\omega t)\,dt.$$

By (4·1) and (4·2), since $k > 0$,

$$\omega\gamma(\omega\eta) \to 0 \text{ as } \omega \to \infty.$$

Hence another form of the necessary and sufficient condition of summability to s at x is

$$\lim_{\eta\to 0}\ \overline{\lim_{\omega\to\infty}}\ \omega^2 \int_0^\eta \Phi(t)\,\gamma'(\omega t)\,dt \Big| = 0 \quad (k > 0). \tag{5·4}$$

6. For the purposes of investigation of the Fourier series of $f(x)$ we define a function $f_1(x)$ which is periodic in 2π, and coincides with $f(x)$ in the interval $(-\pi, \pi)$. Clearly the two functions have, relative to $(-\pi, \pi)$, the same Fourier series.

We proceed to examine the behaviour of $\sigma_k(\omega)$, the partial Rieszian sum of this series, as $\omega \to \infty$.

Write $$A_n(x) = a_n \cos nx + b_n \sin nx.$$

Then by (2·2) $$\sigma_k(\omega) = \tfrac{1}{2}a_0 + \sum_{1\le n<\omega} \left(1 - \frac{n}{\omega}\right)^k A_n(x).$$

THEOREM II. *If* $k > 0$,

$$\sigma_k(\omega) - \tfrac{1}{2}a_0 = \frac{1}{\pi}\int_0^\infty \psi_1\left(\frac{t}{\omega}\right)\gamma(t)\,dt, \tag{6·1}$$

where $$\psi_1(t) = f_1(x+t) + f_1(x-t) - a_0.$$

Then, since $$\frac{1}{\pi}\int_0^\infty \gamma(t)\,dt = \tfrac{1}{2},$$

$\sigma_k(\omega)$ will have the limit s, for $k > 0$, if, and only if,

$$\int_0^\infty \phi_1\left(\frac{t}{\omega}\right)\gamma(t)\,dt \to 0,$$

or, what is equivalent,

$$\omega \int_0^\infty \phi_1(t)\,\gamma(\omega t)\,dt \to 0,$$

where $$\phi_1(t) = f_1(x+t) + f_1(x-t) - 2s.$$

Digitized by Google

7. It is readily verified that $\psi_1(t)$ is periodic in 2π, even, and such that

$$\int_{-\pi}^{\pi} \psi_1(t)\, dt = 0,$$

so that

$$\Psi_1(t) = \int_0^t \psi_1(u)\, du$$

is a continuous function, periodic in 2π.

Now let $v(t)$ be any function of bounded variation in $(0, \infty)$ which tends to zero as $t \to \infty$. Then

$$\begin{aligned}\int_0^\infty \psi_1(t)\, v(t)\, dt &= \lim_{\lambda\to\infty} \int_0^\lambda \psi_1(t)\, v(t)\, dt \\ &= \lim_{\lambda\to\infty} \left\{ \Big[\Psi_1(t)\, v(t) \Big]_0^\lambda - \int_0^\lambda \Psi_1(t)\, dv(t) \right\} \\ &= -\int_0^\infty \Psi_1(t)\, dv(t),\end{aligned}$$

since $\Psi_1(t)$, being continuous and periodic, is bounded in $(0, \infty)$.

If $S_n(t)$ is the nth partial Cesàro sum of the Fourier series of $\Psi_1(t)$,

$$S_n(t) \to \Psi_1(t)$$

uniformly in $(0, \infty)$ by the above property of $\Psi_1(t)$, and it follows immediately that

$$(7{\cdot}1)\quad \lim_{n\to\infty} \int_0^\infty -S_n(t)\, dv(t) = -\int_0^\infty \Psi_1(t)\, dv(t) = \int_0^\infty \psi_1(t)\, v(t)\, dt.$$

The Denjoy-Fourier series of $\psi_1(t)$ proves to be

$$\sum_{n=1}^{\infty} c_n \cos nt$$

where

$$c_n = 2A_n(x),$$

and it may be verified that

$$(7{\cdot}2)\quad \int_0^\infty -S_n(t)\, dv(t) = \sum_{m=1}^{n-1} \left(1 - \frac{m}{n}\right) c_m v_m,$$

where

$$v_m = \lim_{\lambda\to\infty} \int_0^\lambda v(t) \cos mt\, dt.$$

This limit clearly exists.

Digitized by Google

8. By (4·1) and (4·2), $\gamma(t) \to 0$ as $t \to \infty$. It is moreover of bounded variation in $(0, \infty)$ for $k > 0$. Thus we may replace $v(t)$ by $\frac{\omega}{\pi}\gamma(\omega t)$, and the coefficients are found to be

$$\begin{aligned} v_m &= \frac{\omega\Gamma(k+1)}{\pi}\int_0^\infty \frac{C_{k+1}(\omega t)}{(\omega t)^{k+1}}\cos mt\,dt \\ &= \frac{\Gamma(k+1)}{\pi}\int_0^\infty \frac{C_{k+1}(t)}{t^{k+1}}\cos\frac{m}{\omega}t\,dt \\ &= \frac{1}{2}\left(1-\frac{m}{\omega}\right)^k \quad (m \leqslant \omega), \\ &= 0 \qquad\qquad (m \geqslant \omega) \end{aligned}$$

for $k > 0$*.

The series $$\sum_{m=1}^{\infty} c_m v_m$$

is therefore actually finite, and its sum by any positive Cesàro mean coincides with the actual sum. In view of (7·1), for the particular function $v(t)$ we have chosen, (7·2) now becomes

$$\sum_{m=1}^{\infty} c_m v_m = \int_0^\infty \psi_1(t)\, v(t)\, dt$$

or
$$\sum_{1 \leqslant n < \omega}\left(1-\frac{n}{\omega}\right)^k A_n(x) = \frac{\omega}{\pi}\int_0^\infty \psi_1(t)\,\gamma(\omega t)\,dt$$
$$= \frac{1}{\pi}\int_0^\infty \psi_1\left(\frac{t}{\omega}\right)\gamma(t)\,dt.$$

Hence (6·1) holds if $k > 0$, and Theorem II is proved.

The result established may be expressed as follows.

For $k > 0$, the necessary and sufficient condition that the Denjoy-Fourier series (1·1) should be summable (C, k) to s at a point x in $(-\pi, \pi)$ is that

$$\omega\int_0^\infty \phi_1(t)\,\gamma(\omega t)\,dt \to 0 \tag{8·1}$$

as $\omega \to \infty$.

9. Write (8·1) in the equivalent form

$$\frac{\omega}{\pi}\int_0^\infty \psi_1(t)\,\gamma(\omega t)\,dt \to s - \tfrac{1}{2}a_0,$$

or, using (7·1), since $\gamma(t)$ is of bounded variation in $(0, \infty)$ when $k > 0$,

$$-\frac{\omega^2}{\pi}\int_0^\infty \Psi_1(t)\,\gamma'(\omega t)\,dt \to s - \tfrac{1}{2}a_0 \qquad (k > 0).$$

* W. H. Young, *Quart. Journal of Maths.*, 43 (1912), 166.

Digitized by Google

If η is any fixed positive number, $\omega > \dfrac{1}{\eta}$, and M is the upper bound of $|\Psi_1(t)|$,

$$\left|\omega^2 \int_\eta^\infty \Psi_1(t)\,\gamma'(\omega t)\,dt\right| \leqslant M\omega^2 \int_\eta^\infty \frac{A}{(\omega t)^{\alpha+1}}\,dt \quad (\alpha = \min\{k, 2\})$$

$$\leqq \frac{MA}{\omega^{\alpha-1}} \int_\eta^\infty \frac{dt}{t^{\alpha+1}}$$

$$\to 0 \text{ when } k > 1.$$

If $k = 1$, the above shews that we may choose X such that

$$\left|\omega^2 \int_X^\infty \Psi_1(t)\,\gamma'(\omega t)\,dt\right| < \epsilon,$$

where ϵ is any previously assigned positive number.

And then, with the argument used in § 5, it may be shewn by Lebesgue's theorem that

$$\left|\omega^2 \int_\eta^X \Psi_1(t)\,\gamma'(\omega t)\,dt\right| \to 0$$

as $\omega \to \infty$.

Thus, when $k \geqslant 1$, the condition (8·1) becomes

$$-\frac{\omega^2}{\pi} \int_0^\eta \Psi_1(t)\,\gamma'(\omega t)\,dt \to s - \tfrac{1}{2}a_0$$

for fixed positive η.

Moreover,

$$\int_0^\infty t\gamma'(t)\,dt = \lim_{\lambda\to\infty} \int_0^\lambda t\gamma'(t)\,dt$$

$$= \lim_{\lambda\to\infty} \left\{\Big[t\gamma(t)\Big]_0^\lambda - \int_0^\lambda \gamma(t)\,dt\right\}$$

$$= -\frac{\pi}{2} \quad (k > 0),$$

by (4·1), (4·2) and (5·2), and the final form of the condition for summability of the series is

$$\lim_{\eta\to 0}\ \overline{\lim_{\omega\to\infty}} \left|\omega^2 \int_0^\eta \Phi_1(t)\,\gamma'(\omega t)\,dt\right| = 0 \quad (k \geqslant 1). \tag{9·1}$$

Now $\phi_1(t)$ coincides with $\phi(t)$ for $x+t, x-t$ in $(-\pi, \pi)$. Thus (9·1) is equivalent to

$$\lim_{\eta\to 0}\ \overline{\lim_{\omega\to\infty}} \left|\omega^2 \int_0^\eta \Phi(t)\,\gamma'(\omega t)\,dt\right| = 0 \quad (k \geqslant 1). \tag{9·2}$$

Digitized by Google

10. It is convenient at this stage to collect the results established. (5·3) and (5·4), (8·1) and (9·2) shew that for $k \geqslant 1$ the necessary and sufficient condition that the Denjoy-Fourier series and the Denjoy-Fourier integral of $f(x)$ shall be summable (C, k) to s at x is

$$\lim_{\eta \to 0} \overline{\lim_{\omega \to \infty}} \left| \omega^2 \int_0^\eta \Phi(t)\, \gamma'(\omega t)\, dt \right| = 0. \tag{10·1}$$

If $0 < k < 1$ the necessary and sufficient condition of summability of the integral is still (10·1), or it may be written in the equivalent form

$$\lim_{\eta \to 0} \overline{\lim_{\omega \to \infty}} \left| \omega \int_0^\eta \phi(t)\, \gamma(\omega t)\, dt \right| = 0, \tag{10·2}$$

whereas the condition for the series can in general take no more simple a form than

$$\lim_{\omega \to \infty} \omega \int_0^\infty \phi_1(t)\, \gamma(\omega t)\, dt = 0. \tag{10·3}$$

That (10·3) cannot always be reduced to (10·2), i.e. that the series and the integral are not summable (C, k) in the same way for $k < 1$, is best shewn by an example.

11. If the series

$$\tfrac{1}{2} a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

is summable (C, k), where $k > 0$, at all points of a set of positive measure, then, since at these points

$$\frac{a_n \cos nx + b_n \sin nx}{n^k} \to 0,$$

it follows that

$$\frac{a_n}{n^k} \to 0, \quad \frac{b_n}{n^k} \to 0$$

as $n \to \infty$ *.

Thus, if $\frac{a_n}{n^k}$, $\frac{b_n}{n^k}$ do not converge to zero,

$$\frac{a_n \cos nx + b_n \sin nx}{n^k}$$

can only converge to zero at points of a set of measure zero. In other words, unless

$$a_n = o(n^k), \quad b_n = o(n^k)$$

the series

$$\tfrac{1}{2} a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

is summable (C, k) almost nowhere in any interval.

* E. W. Hobson, *Theory of functions of a real variable*, Vol. 2, 682.

Digitized by Google

12. Titchmarsh has shewn* that if $\lambda(n)$ be positive and decrease steadily to zero as $n \to \infty$, then however slowly $\lambda(n) \to 0$ there is a function whose Fourier coefficients are such that

$$a_n \neq o\{n\lambda(n)\}, \quad b_n \neq o\{n\lambda(n)\}.$$

To construct such a function we proceed as follows.

Let $T(x)$ vanish outside the interval $(-\pi, \pi)$. In the interval write

$$T(x) = p^2 n_p \lambda(n_p) \sin n_p x \quad \left(\frac{\pi}{p+1} < x \leqslant \frac{\pi}{p}\right),$$

where $p = 1, 2, 3, \ldots$ and n_p is an integer, depending on p, which tends to ∞ so rapidly that $n_p > 2n_{p-1}$ and $\Sigma p^2 \lambda(n_p)$ is convergent,

$$T(0) = 0,$$

$$T(-x) = -T(x).$$

$T(x)$ has a single point of non-summability in $(-\pi, \pi)$, namely the origin. It is Denjoy integrable in $(-\pi, \pi)$.

As Titchmarsh shews,

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} T(x) \sin nx\, dx \neq o\{n\lambda(n)\}.$$

Thus, for any $\mathbf{k}$, $0 < \mathbf{k} < 1$, we can so arrange $\lambda(n)$ as to make

$$b_n \neq o(n^{\mathbf{k}}),$$

and the Denjoy-Fourier series of $T(x)$ will be summable (C, k), $0 \leqslant k \leqslant \mathbf{k}$, almost nowhere in the interval $(-\pi, \pi)$. The origin is an exceptional point where the series vanishes.

On the other hand, if $k > 0$, the Denjoy-Fourier integral of $T(x)$ is summable (C, k) everywhere in the interval. For it is easily seen that, unless $x = 0$, by a classical argument, since $T(x)$ is absolutely integrable in any interval to which the origin is exterior,

$$\lim_{\eta \to 0} \overline{\lim_{\omega \to \infty}} \left| \omega \int_0^{\eta} (T(x+t) + T(x-t) - 2s)\, \gamma(\omega t)\, dt \right| = 0 \quad (k > 0),$$

where $$s = \frac{T(x+0) + T(x-0)}{2},$$

and at the origin the integral is convergent to zero.

* E. C. Titchmarsh, *Proc. Lond. Math. Soc.*, 22 (1924), Records XXV.

Digitized by Google

Wave Propagation in Strings with Continuous and Concentrated Loads. By Mr HAROLD JEFFREYS, St John's College.

[*Received* 28 April, *read* 2 May, 1927.]

1·0. The normal modes of a stretched string, with equal masses attached at equal intervals along it, are well known. But the method of normal coordinates is not often a very convenient way of treating the motion of systems when started off by impulses or displacements at definite points, unless the number of degrees of freedom is very small, and the most useful in practice is the operational method of Heaviside or the equivalent method of Bromwich depending on the use of complex integrals. Apart from its intrinsic interest, the motion of a loaded string disturbed in this way serves to illustrate several other questions.

First, if the length, total mass, and tension of the string are kept constant, but the individual particles and the distance between them are made very small, we approach in the limit the problem of the uniform heavy string. Now the justification of the operational method is complete for systems with a finite number of degrees of freedom. Starting by defining σ^{-1} to mean the operation of integrating with regard to t from 0 to t, it can be shown without appeal to complex integration that any operator is intelligible if, when we regard σ temporarily as a complex variable, the operator is regular near $\sigma^{-1}=0$ or $\sigma=\infty$. It can also be shown that all the operators arising in the discussion of systems with a finite number of degrees of freedom satisfy this condition*. We may remark that σ is not defined explicitly: only powers of σ^{-1} occur in the operational solution. The operational solution can therefore be completely justified for light strings loaded at a finite number of points.

But when the operational method is applied to continuous systems the solution is found to involve such operators as $e^{-\sigma h/c}$, where h is positive and independent of t. This is not regular near $\sigma^{-1}=0$, and a new rule of interpretation is needed. The rule adopted is that

$$e^{-\sigma h/c}\phi(t)=\phi\left(t-\frac{h}{c}\right) \qquad \text{.....(1).}$$

This rule has been considered a symbolical form of Taylor's theorem; but it gives the correct answer when ϕ or some of its derivatives are discontinuous between t and $t-h/c$, when Taylor's theorem is untrue. If it is interpreted by means of complex integrals it is

* *Cambridge Math. Tracts* (in the press).

Digitized by Google

found to be equivalent, in ordinary cases, to Fourier's theorem. If we approach the continuous system as the limit of a finite one we may hope to find out how the exponential arises and obtain some light on the reason for its interpretation.

Second, the nature of the propagation of a disturbance in a string loaded discontinuously proves to be more complicated than in the continuous string, and dispersion occurs.

2·0. Consider a light string under tension, free to execute either transverse or longitudinal vibrations. Particles of mass ρl are attached to it at intervals l. If $y_0, y_1, y_2 \ldots$ are the displacements of the particles, the kinetic energy is given by

$$2T = \rho l (\dot{y}_0^2 + \dot{y}_1^2 + \dot{y}_2^2 + \ldots) \quad \ldots\ldots\ldots\ldots\ldots\ldots(1)$$

and the potential energy by

$$2V = \rho c^2 l^{-1} \{(y_1 - y_0)^2 + (y_2 - y_1)^2 + \ldots\} \quad \ldots\ldots\ldots(2).$$

For transverse vibrations ρc^2 is the tension; for longitudinal ones it is the sum of Hooke's constant and the tension. The equations of motion are of the form

$$\ddot{y}_r = -\frac{c^2}{l^2}(2y_r - y_{r-1} - y_{r+1}) \quad \ldots\ldots\ldots\ldots\ldots(3).$$

If we make l approach zero while ρ and c remain constant, the system approaches the continuous string with line density ρ, and c becomes the velocity of propagation of waves. If further the quantity on the right of (3) tends to a definite limit we put

$$rl = x \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4),$$

and the equation reduces to the ordinary form

$$\frac{\partial^2 y}{\partial t^2} = c^2 \frac{\partial^2 y}{\partial x^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots(5).$$

Suppose now that the system starts from rest, that the particle corresponding to y_m is kept fixed, and that y_0 is made to vary with the time in some prescribed manner. By our rules we replace the general equation of motion by

$$-\frac{\sigma^2 l^2}{c^2} y_r = 2y_r - y_{r-1} - y_{r+1} \quad \ldots\ldots\ldots\ldots(6),$$

subject to y_0 being given, so that the first equation of type (6) to arise involves $\sigma^2 y_1$. Also

$$y_m = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(7).$$

In obtaining the operational solution of these equations we must treat σ as a constant and carry through the operations as for equations of finite differences. We find that if

$$y_r = A a^r + B b^r \quad \ldots\ldots\ldots\ldots\ldots\ldots(8),$$

Digitized by Google

all the equations for $r = 1$ to $m - 1$ become identities provided that

$$a = e^{\lambda}; \quad b = e^{-\lambda} \quad \text{.....(9),}$$

where

$$\cosh \lambda = 1 + \sigma^2 l^2/2c^2 \quad \text{.....(10),}$$

or

$$\sinh \tfrac{1}{2}\lambda = \frac{\sigma l}{2c} \quad \text{.....(11).}$$

The conditions for $r = 0$ and $r = m$ determine A and B, and our operational solution is easily found to be

$$y_r = \frac{\sinh (m-r)\lambda}{\sinh m\lambda}\, y_0 \quad \text{.....(12).}$$

Now $\sinh m\lambda/\sinh \lambda$ is a polynomial in $\sinh \frac{1}{2}\lambda$ of degree $2(m-1)$, the term of highest degree being $(2 \sinh \frac{1}{2}\lambda)^{2(m-1)}$ or $(\sigma l/c)^{2(m-1)}$. As was to be expected, the operator is therefore a regular function of σ^{-1} near $\sigma^{-1} = 0$, and its first term is $(\sigma l/c)^{-2r}$ or $(c/l\sigma)^{2r}$. It follows at once that the further a particle is from the disturbed end the more gradually it will begin to move.

If we take y_0 to be unity for all positive values of the time the first term in y_r is $\frac{1}{(2r)!}\left(\frac{ct}{l}\right)^{2r}$. If r is great we can approximate to this by Stirling's formula and find

$$y_r \doteqdot \left(\frac{1}{4\pi r}\right)^{\frac{1}{2}} \left(\frac{e}{2}\frac{ct}{rl}\right)^{2r}$$

$$= \left(\frac{l}{4\pi x}\right)^{\frac{1}{2}} \left(\frac{e}{2}\frac{ct}{x}\right)^{2x/l} \quad \text{.....(13).}$$

If then x is greater than $\frac{1}{2}ect$, y_r will tend to zero when l is made to tend to zero. Similar considerations will apply to the later terms in the expansion of the operator, and the fact that in a continuous string it takes a finite time for a disturbance at $x = 0$ to produce any motion at all at a given distance from the end therefore becomes intelligible. This proposition is untrue for the string carrying discrete particles.

Still taking y_0 to be unity when t is positive, and using the complex integral interpretation, we have

$$y_r = \frac{1}{2\pi\iota}\int_L \frac{\sinh (m-r)\lambda}{\sinh m\lambda}\, e^{\gamma t}\, \frac{d\gamma}{\gamma} \quad \text{.....(14),}$$

where the integral is to be taken along a line parallel to the imaginary axis and on the positive side of it, and λ is defined now by

$$\sinh \tfrac{1}{2}\lambda = \frac{\gamma l}{2c} \quad \text{.....(15),}$$

Digitized by Google

subject to λ and γ being real and positive at the same time. Then on the path of integration

$$|e^{\lambda}| > 1 \quad \text{.....(16),}$$

and we can expand in powers of $e^{-2m\lambda}$. The wave-expansion method that occurs so frequently in the operational treatment of continuous systems is therefore applicable to the string with discrete particles. The first term is

$$z_r = \frac{1}{2\pi\iota}\int_L e^{\gamma t - r\lambda}\frac{d\gamma}{\gamma} \quad \text{.....(17),}$$

and the rest follow at once. The first term gives the direct wave, the others the waves reflected at the ends. In operational form we can write

$$z_r = e^{-r\lambda} \quad \text{.....(18)}$$

$$= \left\{\frac{\sigma l}{2c} + \left(\frac{\sigma^2 l^2}{4c^2} + 1\right)^{\frac{1}{2}}\right\}^{-2r} \quad \text{.....(19).}$$

We can evaluate the integral (17) by the method of steepest descents. For given values of t and r there is a saddle point where

$$t = r\frac{d\lambda}{d\gamma},$$

that is, $$\cosh\tfrac{1}{2}\lambda = rl/ct = x/ct = \xi \text{ say} \quad \text{.....(20).}$$

If $\xi > 1$, this makes λ, and therefore γ, real and positive. If $\xi < 1$, there are two saddle points on the imaginary axis, which both contribute to the integral.

If $x > ct$, γ at the saddle point is equal to $(2c/l)\sqrt{(\xi^2 - 1)}$. Near this point

$$\gamma t - r\lambda = \frac{2ct}{l}\left[(\xi^2-1)^{\frac{1}{2}} - \xi\cosh^{-1}\xi\right] + \frac{lt}{4c}\frac{(\xi^2-1)^{\frac{1}{2}}}{\xi^2}(\gamma-\gamma_0)^2 \quad (21),$$

and the line of steepest descent is there perpendicular to the real axis. We find by the usual formula

$$z_r \sim \frac{1}{2}\left(\frac{l}{\pi ct}\right)^{\frac{1}{2}}\frac{\xi}{(\xi^2-1)^{\frac{1}{4}}}\exp\left[-\frac{2ct}{l}\{\xi\cosh^{-1}\xi - (\xi^2-1)^{\frac{1}{2}}\}\right] \quad (22).$$

When ξ is great this approximates to (13).

If $x < ct$, λ at the saddle points will be $\pm i\mu$, where

$$\cos\tfrac{1}{2}\mu = \xi; \quad \gamma = \pm\frac{2c}{l}i\sin\tfrac{1}{2}\mu \quad \text{.....(23).}$$

But it is found that the line of steepest descent through $+i\mu$ is in the direction $\frac{1}{4}\pi$ and goes to $-\infty$; that through $-i\mu$ is in the direction $\frac{3}{4}\pi$ and also goes to $-\infty$. These lines are therefore not

Digitized by Google

together equivalent to the line L, because they pass the origin, which is a pole, on the negative real side. A loop from $-\infty$ around the origin must be added. This makes a contribution unity, and we have

$$z_r \sim 1 - \left(\frac{l}{\pi ct}\right)^{\frac{1}{2}} \frac{\xi}{(1-\xi^2)^{\frac{3}{4}}} \cos\left[\frac{2ct}{l}\{(1-\xi^2) - \xi\cos^{-1}\xi\} + \tfrac{1}{4}\pi\right] \quad (24).$$

Considering now how (22) and (24) behave when l is made small, we see that in front of the point where $x = ct$ the displacement falls off rapidly, and that at all such points it tends to zero with l. Behind this point the displacement is unity except for a rapidly alternating portion, whose wave-length and amplitude both tend to zero with l. The solution for the continuous string is therefore zero displacement for $x > ct$ and unit displacement when $x < ct$. This corresponds to the function $H(t - x/c)$, which is the solution obtained by direct operational methods.

The part of the solution we have just estimated is

$$e^{-r\lambda} H(t) = \left\{\left(1 + \frac{\sigma^2 l^2}{4c^2}\right)^{\frac{1}{2}} + \frac{\sigma l}{2c}\right\}^{-2r} H(t) \quad \ldots\ldots\ldots(25).$$

The operator is regular near $\sigma^{-1} = 0$. If, however, we make l tend to zero, the operator tends (formally) to $\exp(-\sigma rl/c)$ or $\exp(-\sigma x/c)$. The rule that

$$\exp(-\sigma x/c)\, H(t) = H(t - x/c) \quad \ldots\ldots\ldots(26),$$

is therefore justified provided that we define $\exp(-\sigma x/c)$ as meaning

$$\operatorname*{Lim}_{l \to 0} \left\{\left(1 + \frac{\sigma^2 l^2}{4c^2}\right)^{\frac{1}{2}} + \frac{\sigma l}{2c}\right\}^{-2x/l} \quad \ldots\ldots\ldots(27).$$

In this way the exponential operator is defined in terms of the fundamental concept of definite integration. Further, since any function of t can be built up out of the function $H(t)$, this result is immediately generalized to any function.

3·0. Consider next a string whose fixed ends correspond to $r = m$ and $r = -m'$. Initially y and $\dot{y}$ are zero for all particles, except that $y_0 = u$ and $\dot{y}_0 = v$. For positive values of r we still have

$$y_r = \frac{\sinh(m - r)\lambda}{\sinh m\lambda}\, y_0 \quad \ldots\ldots\ldots(1),$$

and for negative values, if $r = -r'$,

$$y_{-r'} = \frac{\sinh(m' - r')\lambda}{\sinh m'\lambda}\, y_0 \quad \ldots\ldots\ldots(2).$$

The subsidiary equation for y_0 is now

$$\left(2 + \frac{\sigma^2 l^2}{c^2}\right) y_0 - y_1 - y_{-1} = \frac{l^2}{c^2}(\sigma^2 u + \sigma v) \quad \ldots\ldots\ldots(3),$$

Digitized by Google

whence, with the same meaning of λ as in the last section,

$$y_0 = \frac{l^2}{c^2}(\sigma^2 u + \sigma v)\frac{\sinh m\lambda \sinh m'\lambda}{\sinh (m+m')\lambda \sinh \lambda} \quad \text{.....(4),}$$

$$y_r = \frac{l^2}{c^2}(\sigma^2 u + \sigma v)\frac{\sinh m'\lambda \sinh (m-r)\lambda}{\sinh (m+m')\lambda \sinh \lambda} \quad \text{.....(5).}$$

If we expand y_r by the wave-expansion method, we find that the first wave is given by

$$z_r = \frac{l^2}{2c^2}(\sigma^2 u + \sigma v)\, e^{-r\lambda}/\sinh \lambda \quad \text{.....(6).}$$

If the string is infinite in length, with the same masses and spacing of the particles, the reflected waves do not arise, and this expression is the complete solution.

3·1. The problem just solved operationally is not of immediate interest, but has important extensions. Suppose first that $y_{-r'}$ and $\dot{y}_{-r'}$ are initially equal to u and v for all values of r'. Then each initial disturbance makes its contribution to y_r and z_r, and we have

$$z_r = \frac{l^2}{c^2}\frac{\sigma^2 u + \sigma v}{2\sinh \lambda} e^{-r\lambda}(1 + e^{-\lambda} + e^{-2\lambda} + \ldots)$$

$$= \frac{l^2}{c^2}\frac{(\sigma^2 u + \sigma v)\, e^{-r\lambda}}{2\sinh \lambda\,(1 - e^{-\lambda})} \quad \text{.....(7).}$$

But by 2·0 (11) this can be written

$$z_r = \left(u + \frac{v}{\sigma}\right)\frac{e^{-(r-\frac{1}{2})\lambda}}{2\cosh \frac{1}{2}\lambda} \quad \text{.....(8)}$$

$$= \frac{1}{2\pi\iota}\int_L \left(u + \frac{v}{\gamma}\right)\frac{e^{\frac{1}{2}\lambda}}{2\cosh \frac{1}{2}\lambda} e^{\gamma t - r\lambda}\frac{d\gamma}{\gamma} \quad \text{.....(9),}$$

λ being now redefined as in 2·0 (15), and the saddle points are as in 2·0. If $x > ct$, we can obtain the solution from 2·0 (22) by multiplying by the value of $\frac{1}{2}(u + v/\gamma)\, e^{\frac{1}{2}\lambda}$ sech $\frac{1}{2}\lambda$ at the saddle point

$$\gamma = \frac{2c}{l}(\xi^2 - 1)^{\frac{1}{2}} \quad \text{.....(10).}$$

The factor required is found to be

$$\tfrac{1}{2}\left\{u + \frac{vl}{2c}\sqrt{(\xi^2 - 1)}\right\}\left\{1 + \frac{(\xi^2 - 1)^{\frac{1}{2}}}{\xi}\right\} \quad \text{.....(11),}$$

and the disturbance when $x > ct$ is again insignificant if l is small. When $x < ct$ the contribution from the saddle points 2·0 (23) is again multiplied by a finite factor and therefore is comparable with

Digitized by Google

its value when y_0 is constrained to be unity. But the contribution from the loop round the origin is altered in form, since the origin is now a double pole. We find that this portion is

$$\tfrac{1}{2}u + \tfrac{1}{2}v\,\{t-(r-\tfrac{1}{2})\,l/c\} \quad \text{.....(12),}$$

or practically

$$\tfrac{1}{2}u + \tfrac{1}{2}v\,\{t-x/c\} = \tfrac{1}{2}\,(u+v/\sigma)\,H\,(-x/c) \quad \text{.....(13).}$$

By making l small and superposing motions due to assigned initial disturbances we can easily obtain the D'Alembert solution for a continuous string

$$y = \tfrac{1}{2}u\,(x+ct) + \tfrac{1}{2}u\,(x-ct) + \frac{1}{2c}\int_{x-ct}^{x+ct} v\,(x)\,dx \quad \text{...(14),}$$

where u and v are now functions of x.

3·2. Let us return now to the case where only particle number 0 is disturbed initially. Using 3·0 (6) we have

$$z_r = \frac{l}{2c}\,\frac{\sigma u + v}{\cosh \tfrac{1}{2}\lambda}\,e^{-r\lambda} \quad \text{.....(15)}$$

$$= \frac{l}{2c}\,\frac{\sigma u + v}{\left\{\dfrac{\sigma l}{2c} + \left(\dfrac{\sigma^2 l^2}{4c^2}+1\right)^{\frac{1}{2}}\right\}^r \left(\dfrac{\sigma^2 l^2}{4c^2}+1\right)^{\frac{1}{2}}} \quad \text{.....(16).}$$

Since the operational form of the Bessel functions is

$$J_n\,(kt) = \frac{\sigma/k}{\left\{\dfrac{\sigma}{k} + \left(\dfrac{\sigma^2}{k^2}+1\right)^{\frac{1}{2}}\right\}^n \left(\dfrac{\sigma^2}{k^2}+1\right)^{\frac{1}{2}}} \quad \text{.....(17),}$$

$z_r - z_{r-1}$ and $\dot{z}_r$ can be expressed explicitly in terms of Bessel functions of integral order and their derivatives, as has been found otherwise by Schrödinger*.

Consider the case where $x = rl < ct$, and interpret (15) as a complex integral. Near a saddle point we have

$$\gamma t - r\lambda = \frac{2ict}{l}\,\{\sqrt{(1-\xi^2)} - \xi\cos^{-1}\xi\} + \frac{1}{4}\frac{ilt}{c}\frac{(1-\xi^2)^{\frac{1}{2}}}{\xi^2}\,(\gamma-\gamma_0)^2 \quad (18),$$

and we find the approximation

$$z_r = \frac{lv}{2c} - \left(\frac{l}{\pi ct}\right)^{\frac{1}{2}}(1-\xi^2)^{-\frac{1}{4}}\left[u\sin\left\{\frac{2ct}{l}\,(\sqrt{(1-\xi^2)} - \xi\cos^{-1}\xi) + \tfrac{1}{4}\pi\right\}\right.$$

$$\left. + \frac{lv}{2c\sqrt{(1-\xi^2)}}\cos\left\{\frac{2ct}{l}\,(\sqrt{(1-\xi^2)} - \xi\cos^{-1}\xi) + \tfrac{1}{4}\pi\right\}\right] \quad \text{...(19).}$$

* *Ann. d. Phys.* 44, 1914, 916–934.

Digitized by Google

When ξ is small, the argument of the periodic terms is nearly $\frac{2ct}{l}(1-\frac{1}{2}\pi\xi)$, the variable part of which is $-r\pi$. Thus the displacements of consecutive particles differ in phase by π. As ξ increases the difference of phase between consecutive particles diminishes, and tends to zero as $\xi \to 1$. The amplitude at the same time increases steadily. But when ξ is nearly unity the argument of the periodic term is nearly $\frac{4}{3}\sqrt{2}\,(1-\xi)^{\frac{3}{2}}\frac{ct}{l}$, and the difference of phase between consecutive particles is comparable with $(1-\xi)^{\frac{1}{2}}$. Also the actual displacements vary only slowly from one particle to the next. It can be shown that analogous relations hold for the velocities.

To estimate $\dot{z}_r$ and $z_r - z_{r-1}$ explicitly the easiest way is to return to the expressions for them as complex integrals. For the former we find

$$\dot{z}_r = \left(\frac{l}{\pi ct}\right)^{\frac{1}{2}}(1-\xi^2)^{-\frac{1}{4}}\left[v\sin\left\{\frac{2ct}{l}(\sqrt{(1-\xi^2)}-\xi\cos^{-1}\xi)+\tfrac{1}{4}\pi\right\}\right.$$
$$\left. + \frac{2c}{l}(1-\xi^2)^{\frac{1}{2}}\,u\cos\left\{\frac{2ct}{l}(\sqrt{(1-\xi^2)}-\xi\cos^{-1}\xi)+\tfrac{1}{4}\pi\right\}\right] \quad (20),$$

and the average kinetic energy of a given particle over a few oscillations is

$$\frac{1}{2}\rho l\,[\dot{z}_r^2]_{\text{mean}} = \frac{1}{4}\frac{\rho l^2}{\pi ct}(1-\xi^2)^{-\frac{1}{2}}\left\{v^2+\frac{4c^2}{l^2}(1-\xi^2)\,u^2\right\} \quad (21).$$

The average potential energy is the same. It can be verified by integration that the total energy is equal to the initial energy.

4·0. It follows easily from the foregoing discussion that the system has no property analogous to heat conduction. Suppose in fact that n consecutive particles are initially given velocities distributed in frequency according to the error law. If the mean square of these velocities is s^2, we can regard s as the initial velocity of agitation of the disturbed region and s^2 as proportional to the temperature. By 3·2 (20) the velocity of agitation s_r of a particle distant rl from the end of the disturbed region is given by

$$s_r^2 = \frac{l}{\pi ct}\,\Sigma\,[v\sin\{tf(\xi)+\tfrac{1}{4}\pi\}]^2(1-\xi^2)^{-\frac{1}{2}} \quad \ldots\ldots\ldots(1),$$

and since by hypothesis the v's are distributed at random they are not correlated with the harmonic factors, and we have

$$s_r^2 = \frac{nl}{2\pi ct}(1-\xi^2)^{-\frac{1}{2}}\,s^2 \quad \ldots\ldots\ldots\ldots\ldots(2),$$

Digitized by Google

where ξ is to be given a value between rl/ct and $(r+n)\,l/ct$. If r/n is large we can ignore the difference and say that (2) gives the distribution of temperature after time t.

This expression bears no resemblance to the solution of the corresponding problem in heat-conduction. If in one-dimensional flow a short stretch of the region is heated and the remainder left at uniform temperature, the temperature at distance x from the heated place after time t is proportional to $t^{-\frac{1}{2}}\exp(-x^2/4h^2t)$, where h^2 is the thermometric conductivity. The factor $t^{-\frac{1}{2}}$ instead of t^{-1} alone is enough to indicate a fundamental difference: the extent of the region containing a given fraction of the energy originally supplied increases in proportion to t in our present problem, instead of $t^{\frac{1}{2}}$ as in the thermal case*.

5·0. In consequence of the failure to find any phenomenon resembling conduction in a string loaded with equal particles equally spaced, it is desirable to investigate whether it can occur when the particles are unequal and their masses are distributed at random. The restriction that they are equally spaced is retained. An operational solution can be obtained without much difficulty, apart from the writing down of large determinants and their expansion in series. Unfortunately, however, no satisfactory way of evaluating the resulting solution has been found. The nature of the solution can be seen more simply by considering a simpler system. Suppose that y_0 is given, that the string is infinite in the positive direction, and that all the particles have mass ρl except one, which has mass $\rho l\,(1+a_k)$. Let its displacement be y_k. Then the waves reaching it are given by

$$y_r = e^{-r\lambda}y_0 \quad \text{.....(1),}$$

and the effect of the exceptional mass is to reflect part of the disturbance. We assume therefore

$$y_r = e^{-r\lambda}y_0 + Ae^{-(k-r)\lambda} \qquad r \leqslant k \quad \text{.....(2),}$$

$$y_r = e^{-(r-k)\lambda}y_k \qquad r \geqslant k \quad \text{.....(3).}$$

But y_r must reduce to y_k when $r=k$, and therefore

$$A = y_k - e^{-k\lambda}y_0 \quad \text{.....(4).}$$

The equation of motion for this particle is

$$-y_{k-1} + \left\{2 + \frac{\sigma^2 l^2}{c^2}(1+a_k)\right\} y_k - y_{k+1} = 0 \quad \text{.....(5),}$$

whence, with

$$\frac{\sigma l}{2c} = \sinh\tfrac{1}{2}\lambda \quad \text{.....(6),}$$

* A general discussion of the related problem of thermal conduction in solids, with references, is given by M. Born, *Atomtheorie des festen Zustandes*, 1923, 708.

Digitized by Google

we find $$y_k = \frac{e^{-k\lambda}}{1 + a_k \tanh \frac{1}{2}\lambda} y_0 \quad \text{.....(7),}$$

and $$y_r = e^{-r\lambda} y_0 - \frac{a_k \tanh \frac{1}{2}\lambda}{1 + a_k \tanh \frac{1}{2}\lambda} e^{-(2k-r)\lambda} y_0 \qquad r \leqslant k \text{ .....(8),}$$

$$y_r = \frac{1}{1 + a_k \tanh \frac{1}{2}\lambda} e^{-r\lambda} y_0 \qquad r \geqslant k \text{ .....(9).}$$

Considering (9) now, we see that k does not appear explicitly. The effect of a particle of abnormal mass upon the transmitted wave is the same whichever it may be. Also, if y_r is found by the method of steepest descents, the situation of the saddle points is unaffected, and we can estimate the transmitted wave at once. If there is a pole at the origin, where λ is zero, its contribution is therefore unaffected. This was to be expected, because in the limit when the particles become indefinitely numerous this part of the solution reduces to the solution for a continuous string, which is obviously unaffected by increasing the mass of an indefinitely short piece by a finite fraction of itself. But the saddle points are at

$$\cosh \tfrac{1}{2}\lambda = \xi = rl/ct \text{ .....(10),}$$

and we must introduce a factor $\left\{1 + ia_k \frac{\sqrt{(1-\xi^2)}}{\xi}\right\}^{-1}$. This factor is complex and therefore gives a shift of phase as well as a change of amplitude. The amplitude is multiplied by

$$\left\{1 + a_k^2 \frac{1-\xi^2}{\xi^2}\right\}^{-\frac{1}{2}} \text{ .....(11).}$$

When ξ is nearly unity, corresponding to the arrival of the first pulse, the amplitude is therefore hardly affected. But as time goes on and ξ decreases, this factor also decreases, and when ξ is small it approximates to ξ/a_k. The energy is at the same time multiplied by $(\xi/a_k)^2$.

This result shows that irregular distribution of mass must have a very marked effect on the propagation of the waves produced by a given initial disturbance. We have here considered such irregularity in its mildest form, where only one particle has a mass differing from the average, and we have found that it reflects the shortest waves almost completely. If many particles have abnormal masses, the small transmission factor will be raised to a high power. A certain fraction of the energy will therefore be concentrated near the region disturbed originally, and will not spread out according to the laws of dispersion. The longest waves, on the other hand, including those that travel fastest, are not much affected by the irregularity. We may say that the effect of irregularity will be to

Digitized by Google

destroy the tail of the train of waves spreading out, and to retain its energy in the form of irregular movement near the origin of the disturbance. Irregularity in the coefficients of $(y_r - y_{r-1})^2$ in the potential energy has also been found to give internal reflexion of similar type.

The effect is very marked. If for instance $\sinh \frac{1}{2}\lambda = \frac{1}{2}i$, corresponding to $\lambda = \frac{1}{3}\pi i$ and a wave length $6l$, ξ is $\frac{1}{2}\sqrt{3}$ and the group velocity is $0{\cdot}866\,c$. But if the irregularity is as great as in an ordinary glass we may suppose a_k^2 on an average $> \frac{1}{100}$ and the transmission by a single particle is under $(1\frac{1}{300})^{-1}$. If then the atomic spacing is 10^{-8} cm. the energy is reduced as the wave travels 1 cm. to $e^{-3{\cdot}3\times 10^4}$ of its initial value. Yet a long wave travels with velocity c without loss. Most of the energy supplied in such a case must therefore be regarded as converted into thermal agitation by internal reflexion.

Summary.

The motion of a light string loaded with equal masses at regular intervals has been discussed by operational methods. It is found that the system, though possessing only a finite number of degrees of freedom, shows most of the characteristic features of dispersion. The operator $e^{-\sigma x/c}$ that occurs in the discussion of uniform continuous strings is found to arise as the limit of an operator defined wholly in terms of definite integration, and its interpretation involves a theorem analogous to Taylor's theorem, but apparently more general. There is no phenomenon analogous to the conduction of heat. Any irregularity in the distribution of mass, however, produces strong internal reflexion of the shortest waves, and may provide a mechanism for conduction.

Digitized by Google

On the Fifth Book of Euclid's Elements. (Second Addendum to Fifth Paper*.) Proof of Euclid v, 19. By Dr M. J. M. HILL, Peterhouse.

[*Received* 5 March, *read* 14 March, 1927.]

There are in the Fifth Book of Euclid's Elements twelve propositions, the object of which is to prove the equality of two ratios.

Euclid deduces five of these, Nos. 4, 7, 11, 12 and 17, directly from the Fifth Definition *without using any other property of ratio.* He uses the 12th proposition in proving the 15th, but a slight change in the form of the proof will make it independent of the 12th.

Simson has given a proof of the same kind of the 18th proposition. This form of proof seems to me the standard form of proof for all propositions, the object of which is to show that some two ratios are equal.

But Euclid's proofs of the remaining five propositions of this group of twelve, Nos. 16, 19, 22, 23 and 24, are of a very different character. They depend on the Seventh Definition (the Test for distinguishing the greater from the smaller of two *unequal* ratios), on the 8th, 10th and 13th propositions, which are concerned with *unequal* ratios, as well as on other propositions. It is very difficult to know which of these should be chosen at any stage of the proof of a proposition in this group.

When I began the study of this subject in 1897 (my first paper was in Vol. XVI of the *Transactions*) I took Euclid's Fifth Definition in the following form:

$$(A : B) = (C : D)$$

if, r, s being *any* integers *whatever*,

and if $rA > sB$, then must also $rC > sD$ .....(i),

if $rA < sB$, then must also $rC < sD$ .....(ii),

if $rC > sD$, then must also $rA > sB$.....(iii),

if $rC < sD$, then must also $rA < sB$.....(iv).

From these it follows by a reductio ad absurdum

that if $rA = sB$, then must also $rC = sD$ .....(v),

if $rC = sD$, then must also $rA = sB$.....(vi).

So that Euclid's conditions are certainly satisfied.

I was not aware at that time that Stolz had shown in 1885 (*Vorlesungen über allgemeine Arithmetik*, Theil I, p. 87) that if the sets of conditions marked (i) and (ii) hold good, then the set marked (v) would follow as a consequence†.

Stolz's proof is printed in my fourth paper, Vol. XXII, p. 188, and also a much simpler proof on p. 189, due to Professor A. C. Dixon and Mr E. Budden.

Thus the two sets of conditions marked (i) and (ii) are equivalent to Euclid's Fifth Definition, and may be treated as a *simplified* form of that Definition.

* The first paper of this set is in the *Transactions*, Vol. XVI, the second in Vol. XIX, the third, fourth and fifth in Vol. XXII. There is an addendum to the fifth paper in Vol. XXI of the *Proceedings*.

† This proposition was known to De Morgan, see my fourth paper, p. 185.

Digitized by Google

There is also contained in the text of Euc. v, 8, of which it forms the greater part, an important proposition, viz.:

"If A, B, C be magnitudes of the same kind, and if A be greater than B, then integers n, t exist, such that

$$nA > tC > nB."$$

Euclid's proof is long and cumbrous. A short proof is given in my fifth paper, Vol. XXII, p. 462.

It depends on Archimedes' Axiom, and so far as the Fifth Book is concerned, this is the only place in which that Axiom is employed. This proposition does not involve any property of ratio.

This proposition contains all that is effective in the Seventh Definition and in Propositions 8, 10 and 13 *for the proof of* properties of equal ratios; and if it be taken in conjunction with the simplified form of the Fifth Definition, referred to above, it is possible to reduce Propositions 16, 22 and 23 to the standard form without difficulty.

Proposition 24 and its Corollaries can likewise be so reduced, though these are more difficult. (*Transactions*, Vol. XXII, p. 460, and *Proceedings*, Vol. XXI, p. 474.)

This leaves only Proposition 19 out of the group of twelve Propositions over. I have given proofs of it (Vol. XIX, p. 166, and Vol. XXII, p. 94) of which the first is very complicated, and the second though less complicated, is hardly a straightforward proof. The proof given below is a better one.

Euclid's proofs of 19 and 24 depend on other properties of ratio and are not straightforward applications of the Fifth Definition.

It is required to prove that if A, B, C, D are four magnitudes of the same kind, and if

$$(A : B) = (C : D),$$

and if

$$A > C,$$

then

$$(A - C : B - D) = (A : B).$$

Using the simplified form of the Fifth Definition, it is enough to consider the alternatives

(i) $$(A : B) > s/r,$$

and (ii) $$(A : B) < s/r,$$

where r, s are *any* positive integers *whatever*.

It will be supposed that the ratios are incommensurable. If they are commensurable there is no difficulty in proving the proposition.

(1) Suppose that $$(A : B) > s/r,$$

then $$rA > sB,$$

and since $$(A : B) = (C : D),$$

$$rC > sD.$$

It has to be proved that

$$r(A - C) > s(B - D).$$

Since $A > C$, the left-hand side of the inequality is positive, and so it is necessary only to consider the case $B > D$.

Digitized by Google

It is necessary to replace

$$rA > sB$$

by

$$mrA > msB$$

so as to obtain a greater difference between the two sides of the inequality.

Let us choose m so large that this difference exceeds $(k+1)B$, where k is an integer to be chosen presently,

then $$m(rA - sB) > (k+1)B,$$

and $$mrA > (ms + k + 1)B \quad \ldots\ldots\ldots\ldots\ldots\ldots (I).$$

Now it is always possible to choose the integer t so that

$$tB < mrA < (t+1)B \quad \ldots\ldots\ldots\ldots\ldots\ldots (II),$$

since A, B are supposed incommensurable.

Further since $(A : B) = (C : D)$ it follows from (II) that

$$tD < mrC < (t+1)D \quad \ldots\ldots\ldots\ldots\ldots\ldots (III).$$

From (II) and (III)

$$mrA + (t+1)D > tB + mrC,$$

hence $$mr(A - C) > t(B - D) - D \quad \ldots\ldots\ldots\ldots (IV).$$

From (II) and (I) it follows that

$$(t+1)B > mrA > (ms + k + 1)B,$$

hence $$t > (ms + k).$$

Using this value of t in (IV)

$$mr(A - C) > (ms + k)(B - D) - D,$$

hence $$mr(A - C) > ms(B - D) + [k(B - D) - D].$$

Now k may be chosen as large as we please.

Let it be chosen so that

$$k(B - D) > D \quad \ldots\ldots\ldots\ldots\ldots\ldots (V).$$

Then it follows that

$$mr(A - C) > ms(B - D),$$

hence $$(A - C : B - D) > s/r.$$

The whole of the work is an algebraic consequence of the inequalities I, II, III and V.

The difficulty is of course to construct them.

(2) Suppose that $(A : B) < s/r,$

then $$rA < sB,$$

and consequently $$rC < sD.$$

Digitized by Google

It has to be proved that

$$r(A-C) < s(B-D).$$

In this case I will begin by writing down the inequality corresponding to (V) in the former case.

It is always possible to choose the integers k, m, t so that

$$k(A-C) > C \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(I)},$$

$$m(sB-rA) > (k+1)A \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(II)},$$

$$tA < msB < (t+1)A \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(III)}.$$

Consequently

$$tC < msD < (t+1)C \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(IV)}.$$

The remainder of the work is an algebraic consequence of the four inequalities (I), (II), (III) and (IV).

From (III) and (IV)

$$msB + (t+1)C > tA + msD,$$

hence

$$ms(B-D) > t(A-C) - C \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(V)}.$$

From (II)

$$msB > (mr+k+1)A \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(VI)}.$$

From (III)

$$(t+1)A > msB \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{(VII)}.$$

From (VI) and (VII)

$$(t+1)A > (mr+k+1)A,$$

hence

$$t > mr + k.$$

Using this value of t in (V)

$$ms(B-D) > mr(A-C) + [k(A-C) - C].$$

Then from (I) it follows that

$$ms(B-D) > mr(A-C),$$

hence

$$(A-C : B-D) < s/r.$$

(3) Hence it has been shown that if

$$(A:B) > s/r, \text{ then } (A-C : B-D) > s/r,$$

if

$$(A:B) < s/r, \text{ then } (A-C : B-D) < s/r.$$

Hence s/r, which represents *any* rational fraction *whatever*, cannot lie between the ratios

$$(A:B) \text{ and } (A-C : B-D),$$

therefore

$$(A-C : B-D) = (A:B).$$

This proof is of the same kind as those previously given for Propositions Nos. 16, 22, 23 and 24, it being based on the Simplified form of the Fifth Definition and the earlier part of the 8th proposition referred to above, which earlier part is not concerned with ratios. Archimedes' Axiom is involved. No other properties of Proportion are employed, as in Euclid's proof.

Digitized by Google

Some Measurements on the Absorption of X-Rays of Long Wave-length. By L. H. MARTIN, M.Sc.,* Trinity College. (Communicated by Prof. Sir E. RUTHERFORD.)

[*Read* 14 March 1927.]

§ 1. *Introduction.*

An account is given here of the measurements of the mass absorption coefficients of iron, nickel, copper and aluminium over a wave-length range 0·705 Å.U. to 1·932 Å.U.

As these measurements were made primarily for use in another experiment, described elsewhere, in connection with which a high degree of accuracy was not necessary, they can only be regarded as preliminary. The results, however, show some interesting features which seem to justify their publication.

Measurements in this range of wave-lengths possess a particular interest, as, in virtue of the magnitude of the "total" mass absorption coefficient, the uncertain quantity representing the "scattering" absorption can be neglected in the investigation of the relationship between the "true" or "fluorescent" mass absorption coefficient, and the wave-length of the absorbed radiation. The precise investigation of this relationship has lately acquired an added interest, as the form of the relation deduced in a recent paper by Wentzel† making use of the methods of Schrödinger's wave mechanics, differs somewhat from the well-known forms given by de Broglie‡, deduced semi-empirically, and by Kramers§, deduced from Correspondence Principle considerations.

Measurements have been made previously by Barkla and Sadler‖ for these elements over this range of wave-lengths. As sources of radiation, Barkla and Sadler used the K characteristic radiations excited in a series of radiators extending from silver to chromium, but as the constituents of such beams differ greatly in wave-length, the absorption coefficients determined differ considerably from the values corresponding to the wave-length of the $K\alpha$ radiation.

As the mean absorption coefficient for a heterogeneous beam of X-rays is a function of the thickness of the absorbing material used in their experiments, it is not possible to assign with any certainty "mean" wave-lengths to the series of radiations used.

* These measurements were made during the period of an International Education Board Fellowship held by the writer.

† Wentzel, *Zeit. f. Phys.* 40, 1926, p. 575.

‡ de Broglie, *Journ. de Phys.* 3, 1922, p. 33.

§ Kramers, *Phil. Mag.* 46, 1923, p. 836.

‖ Barkla and Sadler, *Phil. Mag.* 14, 1909, p. 739.

Digitized by Google

Hewlett* carried his measurements for aluminium and iron down to 1·006 Å.U.

In the present experiments, the usual direct method for the determination of absorption coefficients has been employed. Homogeneous beams were obtained in the ordinary manner by crystal reflection, and in the case of the longer wave-lengths, suitable targets were employed in the X-ray tube to give characteristic radiations of high intensity. A method was devised which made it possible to employ the same metal foil for absorption measurements on each side of a K discontinuity.

In the case of Fe, Ni and Cu, it was found that when the X-rays absorbed lie in the range on the short wave-length side of the respective K absorption discontinuities, the mass absorption coefficients are represented with good approximation by a relation of the form

$$\mu/\rho = \text{const. } \lambda^{2\cdot8} \quad \ldots\ldots\ldots\ldots\ldots\ldots (1).$$

On the long wave-length side of the K discontinuities and in the case of aluminium throughout the entire wave-length range the absorption coefficients are represented with good approximation by

$$\mu/\rho = \text{const. } \lambda^{3} \quad \ldots\ldots\ldots\ldots\ldots\ldots (2).$$

Recently S. J. Allen† has published an account of experiments in which he has been able to extend considerably his previous measurements of the absorption coefficients of a number of elements from carbon to uranium from $\lambda = 0{\cdot}709$ Å.U. to 4·0 Å.U. The results of the present paper are later compared with those of Allen, and discussed in relation to the different expressions which have been deduced theoretically relating μ/ρ with the atomic number Z and the wave-length λ of the absorbed radiation.

§ 2. *Experimental details.*

As intense beams of X-rays of long wave-length were required, it was necessary to use an X-ray tube provided with a thin aluminium window. Different types, including gas tubes, were tried, but the one shown in Fig. 1 was found to be the most satisfactory. This is of the hot cathode type. The electrodes can be withdrawn easily and replaced by the ground-glass joints provided at each end of the tube. The electrodes are surrounded by a brass tube 3 cms. in diameter which is earthed. This protects the glass walls of the X-ray tube from local heating due to scattered electrons, and also has the effect of causing the tube to run "hard." (With the Hyvac alone pumping and the filament off there was no sign of a discharge at 60—75 K.V.) The tube is exhausted continuously

* Hewlett, *Phys. Rev.* 17, 1921, p. 284.
† S. J. Allen, *Phys. Rev.* 28, 1926, p. 907.

Digitized by Google

by a Crawford mercury vapour jet pump (Western Electric Co. pattern) and runs quite steadily when the current is not greater than 5 milliamps. In these experiments the cathode was earthed, and it may be remarked that filaments (5 amp. tungsten wire) then last six months or longer, even with practically continuous running.

A measure of the intensity of the X-ray beam issuing from the tube was obtained by placing a small ionisation chamber between the collimator slits of the spectrometer, and measuring the ionisation produced by means of a Dolezalek electrometer. Since this ionisation chamber measures the total intensity of the beam, it is important to maintain a constant potential across the X-ray tube in order to keep constant the distribution of relative intensities among the various wave-lengths. To this end a voltmeter was placed across the primary of the transformer, and the primary voltage controlled by means of rheostats in series with the primary windings. The tungsten filament was heated from accumulators. The spot from the Dolezalek electrometer, and also that from a Compton electrometer connected with the spectrometer

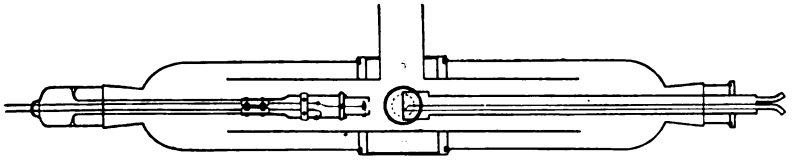


Fig. 1.

ionisation chamber, were arranged to pass across two scales placed one immediately above the other.

For wave-lengths near to and on the short wave-length side of the K absorption edges of the metals used, a special difficulty arises in obtaining foils which are at once thin enough to transmit a useful percentage of the incident radiation, and sufficiently uniform to justify the determination of the mass per unit area from measurements made on a piece of foil of the order 1 sq. inch in area. (The half value thickness for Cu $K\alpha_1\alpha_2$ in iron is approx. 0·0003 cm.)

Foils of this effective thickness were produced by mounting a thicker foil on the moving arm of a divided circuit. The intensity is measured first with the foil perpendicular to the X-ray beam, and then when the foil is turned through a known angle θ. The decrease in intensity is equivalent to that caused by a foil of thickness

$$d = D(\sec\theta - 1),$$

where D is the actual thickness of foil used.

A typical set of readings is shown in Table 1.

Digitized by Google

TABLE 1.

IRON foil. 0·0117 gm. cm.$^{-2}$ $\lambda = 1{\cdot}537$ A.U.
Effective mass per unit area $0{\cdot}0117\,(\sec 30^\circ - 1) = 0{\cdot}00180$ gm. cm.$^{-2}$.

Intensities		
I_d $\theta = 30^\circ$	I_0 $\theta = 0^\circ$	I_d $\theta = 30^\circ$
2·50	4·31	2·20
2·42	4·24	2·29
2·23	4·10	2·14
2·38	4·13	2·44

This method of realising thin foils can be of use only when the initial radiations are of high intensity. In the present experiment, intense homogeneous beams were obtained by reflecting from a calcite crystal the characteristic radiations emitted from different targets placed on the anticathode of the X-ray tube. Quite wide slits may then be used, as the intensity of the general radiation is negligible compared with that of the characteristic radiations. The potential of the tube was controlled so that no second order radiation was reflected. In the case of copper, it was not possible to obtain measurements near the K limit on the short wave-length side, as the intensity of the general radiation from the platinum target excited at the low potentials necessary to prevent second order reflection was too small to permit of the method described above for obtaining thin foils. Foils of the required thickness could not be obtained sufficiently uniform.

In order to eliminate as far as possible lack of uniformity in the metal foils, small foils approximately one inch square were used in the determination of the mass per unit area, and in the absorption experiments, different sections of the absorber were placed in turn in the path of the X-ray beam.

The accuracy of the following results is determined by the degree of steadiness which could be maintained in the tube, but principally by the extent of the elimination of lack of uniformity of the foils. In most cases the error is estimated at 3 %, but may be greater for values near to and on the short wave-length side of a K discontinuity.

§ 3. *Results.*

The results are collected in Fig. 2, where $\log \mu/\rho$ is plotted against $\log \lambda$ and from which the following relations are deduced.

The only results available for comparison are the recent measurements of Allen, and the much earlier ones of Barkla and Sadler. Table 3 shows the different values of μ/ρ found for the $K\alpha_1\alpha_2$ radiations of the targets Fe, Ni, Cu and Zn.

Digitized by Google

ALUMINIUM foils. Mass per unit area 0·1222 gm. cm.$^{-2}$
" " " " 0·09152 gm. cm.$^{-2}$.
Values corrected for 0·31 °/$_{\circ}$ iron content.

Wave-length Å.U.	Slit width mm.	μ/ρ
0·705	·15	4·98
0·799	·15	7·17
0·965	·15	12·1
1·100	·15	18·2
1·293	·08	30·3
1·433	·08	40·4
1·537	·08	51·4
1·655	·08	63·9
1·932	·08	91·8

IRON foil, obtained by electrolytic deposition.
Mass per unit area 0·0117 gm. cm.$^{-2}$

Wave-length Å.U.	Slit width mm.	μ/ρ
0·705	·15	36·7
0·799	·15	53·1
0·965	·15	87·1
1·100	·20	130
1·293	·35	200
1·433	·30	265
1·537	·30	316
1·655	·30	416
1·932	·08	68·8

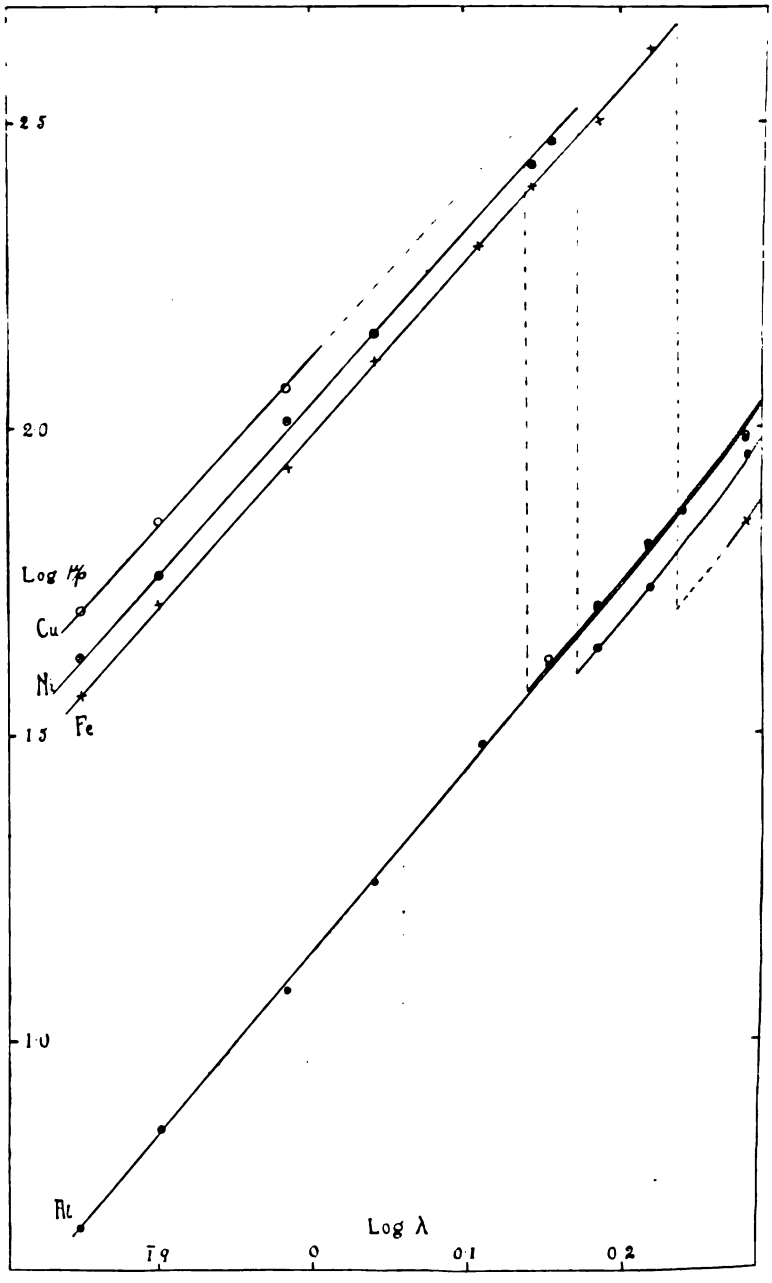
NICKEL foil, obtained by electrolytic deposition.
Mass per unit area 0·01362 gm. cm.$^{-2}$
" " " " 0·01142 gm. cm.$^{-2}$

Wave-length Å.U.	Slit width mm.	μ/ρ
0·705	·15	42·1
0·799	·15	58·8
0·965	·15	103
1·100	·20	145
1·389	·30	272
1·432	·08	298
1·537	·08	43·9
1·655	·08	53·8
1·932	·08	87·8

COPPER foils. Mass per unit area 0·01508 gm. cm.$^{-2}$
" " " " 0·01252 gm. cm.$^{-2}$

Wave-length Å.U.	Slit width mm.	μ/ρ
0·705	·15	49·6
0·799	·15	72·4
0·965	·15	118
1·432	·08	41·8
1·537	·08	50·1
1·655	·08	62·4
1·932	·08	95·3

Digitized by Google



F ig. 2.

Digitized by Google

TABLE 2.

Element	·705 Å.U. $\rightarrow \lambda_K$ $\frac{\mu}{\rho}(K+L+\ldots)$	$\lambda_K \rightarrow$ 1·932 Å.U. $\frac{\mu}{\rho}(L+M+\ldots)$
Fe	$98\lambda^{2\cdot80}$	—
Ni	$110\lambda^{2\cdot78}$	$12\cdot4\lambda^{3}$
Cu	$130\lambda^{2\cdot80}$	$14\cdot1\lambda^{3}$
	$\frac{\mu}{\rho}(K+L+\ldots)$	
Al	$13\cdot9\lambda^{2\cdot98}$	

TABLE 3. *Values of* μ/ρ.

Absorber	Al			Fe			Ni			Cu		
Observer	B.&S.	A.	M.	B.&S.	A.	M.	B.&S.	A.	M.	B.&S.	A.	M.
Radiation												
Fe Ka_1a_2	88·5	94·0	91·8	66·1	71	68·8	83·8	90	87·8	95·1	99	95
Ni Ka_1a_2	59·1	61·5	63·9	314	430	416	56·3	60	53·8	61·8	64	62·4
Cu Ka_1a_2	47·7	48·5	51·4	268	330	316	(62·7)	47·5	43·9	53·0	50	50·1
Zn Ka_1a_2	39·4	40·3	40·4	221	290	265	265	323	298	(55·5)	42·0	41·8

Those readings of Barkla and Sadler which are enclosed in brackets refer to cases in which the β_{12} constituents of the K radiations are of shorter wave-length than the K limit of the absorbing material, and therefore lead to values of μ/ρ which are too high.

In the case of aluminium, the agreement shown is quite satisfactory. The smaller values of Barkla and Sadler would be expected as the weak short wave components $K\beta_{12}$ are present in the beams of X-rays used by them. The agreement between the values of Allen and those given here for aluminium persists throughout the whole range of wave-lengths, and indeed, Allen finds that in the wave-length range λ 0·08 Å.U. to λ 4·0 Å.U. the absorption coefficients of aluminium are given with some accuracy by the formula

$$\mu/\rho = (13\cdot95 \pm 0\cdot05)\,\lambda^{2\cdot92} + (0\cdot16 \pm 0\cdot02),$$

a relation to which that given here (Table 2) is equivalent within the limits of experimental error.

Digitized by Google

A satisfactory agreement is found too in the case of Fe (see Table 3) and Cu, but in the case of Ni, the values found here are in general from 5 °/₀ to 10 °/₀ smaller than those given by Allen.

It is difficult to account for this discrepancy, particularly as the general variation of μ/ρ with λ is the same in both experiments. With regard to the present experiments possible errors introduced by lack of uniformity of the foil used or even in the determination of the mass per unit area would be reduced to a minimum by the use of two separate specimens of nickel. As mentioned above, care was taken to prevent any second order radiation being present.

The following values were found for the K jump, i.e. the ratio of the absorption coefficient on the short wave-length side to that on the long wave-length side at the K limit. It must be pointed out that the higher value refers to $\lambda = 0{\cdot}7$ Å.U. and is deduced on the assumption that the L absorption obeys a λ^3 law over the range, $\lambda_K \rightarrow 0{\cdot}7$ Å.U. Allen finds as the result of his more extensive measurements that the K jump for these elements reaches its maximum value in the neighbourhood of $\lambda = 0{\cdot}6$ Å.U., when the exponent of λ for the K absorption approximates to that for the L absorption on the long wave-length side of the K limit.

The smaller value given for the K jump is its value at the K limit.

TABLE 4.

K jumps.

$$\frac{\mu}{\rho}(K+L+\ldots) \Big/ \frac{\mu}{\rho}(L+M+\ldots)$$

Element	Allen	Martin
Fe	9·5—10	9·2—10
Ni	8·8— 9·8	8·5— 9·6
Cu	9·1— 9·8	8·6—10

§ 4. *Discussion of results.*

Several expressions have been deduced theoretically relating the "fluorescent" absorption coefficient τ with the fundamental quantities Z, the atomic number of the absorbing element, and λ, the wave-length of the absorbed radiation.

De Broglie assumes that the laws of black body radiation can be extended to the case of X-ray absorption, and deduces semi-empirically the following expression for the atomic absorption coefficient

$$\tau_A = \text{const. } \lambda^3 \sum_n \frac{b_n}{\lambda_n^3} \quad \ldots\ldots\ldots\ldots\ldots\ldots (1),$$

where b_n is the number of electrons in the n quantum group and λ_n is the wave-length of the corresponding absorption discontinuity.

Digitized by Google

Kramers, by means of the application of the Correspondence Principle to a "hydrogen-like" atom, deduced the following expression

$$\tau_A = \text{const. } \lambda^3 Z^4 \Sigma_n \frac{b_n}{a_n n^3} \quad \ldots\ldots\ldots\ldots\ldots\ldots(2),$$

where a_n is the statistical weight of the electrons in the n quantum group.

In both (1) and (2) the summation is to be extended only over those groups of electrons whose critical frequency is less than that of the radiation considered. Recently Bothe* has considered the effect of those electron groups which correspond to critical frequencies which are greater than that of the absorbed radiation. He finds the following relation for the atomic absorption coefficient when $\lambda \ll \lambda_n$

$$\mu_A = \frac{2\pi\rho^2}{mc^2} \lambda^3 \Sigma_n \frac{b_n}{\lambda_n^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots(3).$$

When $\lambda > \lambda_n$ the ratio of the absorption coefficient μ_n to that at λ_n on the short wave-length side of the discontinuity, μ_n^0, is given by

$$\frac{\mu_n}{\mu_n^0} = \frac{3{\cdot}74 \times 10^{-5}}{\lambda_n} \left[\frac{2r^2 - 1}{r^2 - 1} - 2r^2 \log \frac{r^2}{r^2 - 1} \right],$$

where

$$r = \lambda/\lambda_n.$$

As such an absorption cannot be associated with the ejection of photoelectrons, Bothe concludes that at any discontinuity a large increase in the scattering coefficient should be found on the long wave-length side.

In the present experiments the scattering coefficient is so small compared with the fluorescent coefficient, that even an increase in the scattering to several times its normal value would have escaped detection.

Recently Wentzel (*l.c.*) has applied the methods of Schrödinger's wave mechanics for a "hydrogen-like" atom to obtain an expression for the K absorption coefficient which differs considerably from equations (1), (2) and (3), viz.

$$\frac{\mu}{\rho} \mathbf{A} = 1{\cdot}82 \times 10^{12} \left(\frac{\lambda}{\lambda_K} \right)^{5/2} [3(A)\lambda^{1/2} + (B)\lambda_K^{1/2}]^2 \quad \ldots(4),$$

or, replacing λ_K by a constant multiple of Z^{-2},

$$\frac{\mu}{\rho} \mathbf{A} = aZ^5 \lambda^{7/2} + bZ^4 \lambda^3 + cZ^3 \lambda^{5/2} \quad \ldots\ldots\ldots\ldots\ldots(5),$$

* W. Bothe, *Zeit. f. Phys.* 1927, p. 653.

Digitized by Google

[illegible]

[illegible]

[illegible]

A somewhat similar breakdown in the λ^3 law near a K limit has been observed by Richtmeyer in the case of Mo, and by Stoner and Martin in the case of Sn, Ag and Pd. Stoner and Martin found that from $\lambda = 0{\cdot}3$ Å.U. to the K jump, the rate of increase of μ/ρ with wave-length is slightly greater than $\lambda^{5/2}$, while on the long wave-length side of the K limit, the λ^3 law was found to hold with good approximation.

These results, however, are in complete disagreement with some precise measurements made recently by Richtmeyer and Bishop* near the K absorption limit of silver. They find that after correcting for the finite width of the slits, the λ^3 law holds on each side of the K limit.

In the present experiments, however, it must be mentioned that although wide slits have been used near the K limit, no appreciable error can have entered on this account, because the radiations used are homogeneous and of high intensity relative to the general radiation. These experiments show that in the case of iron, nickel and copper, there is a definite departure from the λ^3 law near the K limits.

The extent to which the Z^4 law holds is shown in the following table:

TABLE 6.

Absorber	$\left(\frac{\text{Atomic Nos.}}{26}\right)^4$	Relative values of $\mu/\rho \times$ Atomic weight			
		$\lambda = 1{\cdot}932$ Å.U.		$\lambda = {\cdot}965$ Å.U.	
		Allen	Martin	Allen	Martin
Fe	1	1	1	1	1
Ni	1·35	1·37	1·38	1·27	1·28
Cu	1·55	1·50	1·50	1·43	1·46

Table 6 indicates the existence of a considerable departure from a Z^4 relation on the short wave-length side of the K limit, and it would appear that a generalised expression similar to equation (5) in which the factors a, b and c are functions of Z and λ seems necessary to represent the values of μ/ρ in an extended wave-length range. Wentzel points out that α depends ultimately on λ and Z and it is possible that when, as a result of the development of wave mechanics, α can be evaluated, a form of (5) will be found to fit experimental results.

In conclusion, the author wishes to express his gratitude to Professor Sir Ernest Rutherford for the continued advice and assistance received from him during the course of this experiment.

* Richtmeyer and Bishop, *Phys. Rev.* 27, 1926, p. 294.

Digitized by Google

The total energy of binding of a heavy atom. By Mr E. A. Milne
Trinity College.

[*Received* 4 May, *read* 25 July, 1927.]

1. In a recent important paper*, L. H. Thomas has originated a method for calculating approximately the field in the interior of a heavy atom. His main assumption is that the electrons are distributed uniformly in the six-dimensional phase space for the motion of an electron at the rate of two for each h^3 of 6-volume. On this assumption, Poisson's equation

$$\nabla^2 V = -4\pi\rho$$

is found to lead to an equation of the form

$$\frac{1}{\xi^2}\frac{d}{d\xi}\left(\xi^2\frac{d\phi}{d\xi}\right) = \phi^{\frac{3}{2}}, \quad \ldots\ldots\ldots\ldots\ldots\ldots(1)$$

where ϕ is proportional to the potential V and ξ is proportional to the radial distance r.

To an astrophysicist, the form of (1) immediately suggests that of Emden's† differential equation

$$\frac{1}{\xi^2}\frac{d}{d\xi}\left(\xi^2\frac{d\phi}{d\xi}\right) = -\phi^n, \ldots\ldots\ldots\ldots\ldots\ldots(2)$$

which describes the gravitational equilibrium of a spherical mass of perfect gas of polytropic index n. Here ϕ may be considered to be either the gravitational potential or the temperature. Emden's solution of (2) with $n = 3$ is the basis of Eddington's famous theory of the internal constitution of the stars. The methods used by Thomas for discussing (1) are similar to those used by Emden for discussing (2).

Amongst Emden's results, perhaps the most notable is that which shows that once (2) has been solved, the total gravitational potential energy (or the mean temperature)‡ can be calculated without further quadrature. The object of the present note is to apply Emden's method to determine the total electrostatic potential energy of an atom built on Thomas's model, and so the total binding energy of the electrons.

2. *Evaluation of an integral.* We consider first a generalised form of Thomas's equation,

$$\frac{1}{\xi^2}\frac{d}{d\xi}\left(\xi^2\frac{d\phi}{d\xi}\right) = \phi^n. \quad \ldots\ldots\ldots\ldots\ldots\ldots(3)$$

* "The calculation of atomic fields," *Proc. Camb. Phil. Soc.*, 23, 1927, 542.
† Emden, *Gaskugeln*, Leipzig, 1907.
‡ This adaptation is due to Eddington.

Digitized by Google

From the mode of derivation of the equation, the charge density ρ is proportional to ϕ^n, whilst $V \propto \phi$, and the total electrostatic energy $\frac{1}{2}\int \rho V d\tau$ thus involves the evaluation of a definite integral

$$I_a^b = \int_a^b \phi^{n+1} \xi^2 d\xi.$$

Substituting for $\xi^2 \phi^n$ from (3) and integrating by parts, this gives

$$I_a^b = \left[\xi^2 \phi \frac{d\phi}{d\xi}\right]_a^b - \int_a^b \left(\frac{1}{\xi^2}\right)\left(\xi^3 \frac{d\phi}{d\xi}\right)^2 d\xi,$$

where the new integrand has been re-arranged as a product of two factors. Integrating again by parts and then substituting for

$$\frac{d}{d\xi}\left(\xi^3 \frac{d\phi}{d\xi}\right)$$

from (3), we find

$$I_a^b = \left[\xi^2 \phi \frac{d\phi}{d\xi} + \xi^3 \left(\frac{d\phi}{d\xi}\right)^2\right]_a^b - 2\int_a^b \xi^3 \left(\phi^n \frac{d\phi}{d\xi}\right) d\xi.$$

On integrating the last integral by parts (treating $\phi^n d\phi/d\xi$ as the perfect differential), we obtain an integrated part together with a multiple of the original integral I_a^b. Re-arranging, we find

$$I_a^b = \frac{n+1}{n-5}\left[\xi^2 \phi \frac{d\phi}{d\xi} + \xi^3 \left(\frac{d\phi}{d\xi}\right)^2 - \frac{2}{n+1}\xi^3 \phi^{n+1}\right]_a^b.$$

This is, essentially, the analogue for equation (3) of Emden's result for (2).

Applying to Thomas's case and extending the integration from 0 to ∞ we find

$$\int_0^\infty \phi^{\frac{5}{2}} \xi^2 d\xi = \frac{5}{7} \lim_{\xi \to 0} \xi^2 \frac{d\phi}{d\xi}\left(\phi + \xi \frac{d\phi}{d\xi}\right).$$

For an atomic field, ϕ has a singularity at the origin, its behaviour being of the form

$$\phi \sim \frac{a_0}{\xi} + a_1$$

for ξ small. We find then

$$\int_0^\infty \phi^{\frac{5}{2}} \xi^2 d\xi = -\frac{5}{7} a_0 a_1. \quad \ldots\ldots\ldots\ldots\ldots\ldots (4)$$

3. *The electrostatic potential energy.* Thomas's equation for the atomic field V at distance r is

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2 \frac{dV}{dr}\right) = -4\pi\rho, \quad \ldots\ldots\ldots\ldots\ldots\ldots (5)$$

with

$$\rho = -e \frac{2}{h^3} \frac{4\pi}{3} (2me)^{\frac{3}{2}} V^{\frac{3}{2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (6)$$

Digitized by Google

and the solution V for r small behaves like

$$V \sim \frac{Ne}{r} + v_0,$$

where N is the atomic number and v_0 is the self-potential at the origin of the distribution ρ. Now the system of charges consists of $+Ne$ at $r = 0$ together with the distribution ρ. Let the latter give rise to a field of potential $v(r)$. Then

$$V = \frac{Ne}{r} + v(r).$$

The electrostatic energy is then $W = \frac{1}{2}\Sigma eV$

or
$$W = \tfrac{1}{2}\int \rho V d\tau + \tfrac{1}{2} N e v_0 \ldots\ldots\ldots\ldots\ldots\ldots (7)$$

Thomas makes the substitutions

$$v = \xi \frac{h^2}{4\pi^2 me^2},$$

$$V = \frac{9\pi^2}{128} \frac{4\pi^2 me^2}{h^2} \phi,$$

so that for ξ small, the behaviour of ϕ is

$$\phi \sim \frac{128}{9\pi^2} \frac{N}{\xi} + \omega_0, \ldots\ldots\ldots\ldots\ldots\ldots (8)$$

say. With these substitutions, (5) reduces to (1), the integral in (7) reduces to a multiple of (4), and (7) itself reduces to

$$W = \frac{9\pi^2}{128} \frac{2\pi^2 me^4}{h^2} N \omega_0 \left[\frac{5}{7} + 1\right], \ldots\ldots\ldots\ldots\ldots\ldots (9)$$

where in the [ ] we have shown separately the contributions corresponding to the two terms in (7).

We shall put

$$\chi = \frac{2\pi^2 me^4}{h^2},$$

the ionisation potential of the normal hydrogen atom.

By the theorem of the virial, since the atomic cluster of electrons is in a steady state, the average kinetic energy is equal to one-half the average potential energy taken with negative sign. Consequently the total energy E of the system of moving charged masses —itself a negative quantity—is one-half expression (9), or

$$E = \frac{6}{7} \cdot \frac{9\pi^2}{128} \chi N \omega_0. \ldots\ldots\ldots\ldots\ldots\ldots (10)$$

It remains to evaluate ω_0.

Digitized by Google

4. *Value of* ω_0. Suppose that by any means whatever we have found a solution of (1) which near $\xi = 0$ is of the form

$$\phi \sim \frac{b_0}{\xi} + b_1. \quad \text{.....(11)}$$

Make the substitution $\xi = \alpha\xi_1$, $\phi = \beta\phi_1$. Then

$$\frac{1}{\xi_1^2}\frac{d}{d\xi_1}\left(\xi_1^2\frac{d\phi_1}{d\xi_1}\right) = \alpha^2\beta^{\frac{1}{2}}\phi_1^{\frac{3}{2}}$$

and, near $\xi = 0$,

$$\phi_1 \sim \frac{b_0}{\alpha\beta\xi} + \frac{b_1}{\beta}.$$

It follows that if we choose α and β so that

$$\alpha^2\beta^{\frac{1}{2}} = 1,$$

$$\frac{b_0}{\alpha\beta} = \frac{128}{9\pi^2}N,$$

then ϕ_1 is precisely the ϕ of equation (8), and accordingly

$$\omega_0 = \frac{b_1}{\beta}.$$

Solving for α and β we find

$$\omega_0 = b_1\left(\frac{128}{9\pi^2}\frac{N}{b_0}\right)^{\frac{4}{3}}. \quad \text{.....(12)}$$

Now Thomas has tabulated* a function ψ_0 given by

$$\psi_0 = \frac{9\pi^2}{128}\phi,$$

(where ϕ satisfies (1)) and having the behaviour near $\xi = 0$,

$$\psi_0 \sim \frac{55}{\xi} + c_1.$$

Thus we know a solution in which

$$b_0 = \frac{128}{9\pi^2}.55, \quad b_1 = \frac{128}{9\pi^2}c_1.$$

Hence

$$\omega_0 = \frac{128}{9\pi^2}\left(\frac{N}{55}\right)^{\frac{4}{3}}c_1,$$

and thus by (10)

$$E = \frac{6}{7}\frac{\chi N^{\frac{7}{3}}}{55^{\frac{4}{3}}}\lim_{\xi\to 0}\left(\psi_0 - \frac{55}{\xi}\right). \quad \text{.....(13)}$$

* We have replaced Thomas's ρ throughout by ξ, to avoid confusion with our use ρ for charge-density.

Digitized by Google

5. *Numerical values.* The last five entries in Thomas's table yield values of $\psi_0 - 55/\xi$ as follows:

ξ	ψ_0	$\psi_0 - 55/\xi$
·003811	14140	−290
·003027	17870	−300
·002404	22580	−300
·001910	28500*	−300
·001517	35960	−300

We have a well-defined value $c_1 = \lim(\psi_0 - 55/\xi) = -300$. Inserting this in (13) we get

$$E = -1{\cdot}23\chi N^{\frac{7}{3}} \quad \text{.....(14)}$$

or in volts, putting $\chi = 13{\cdot}54$, and omitting the minus sign,

$$E = 17N^{\frac{7}{3}} \text{ volts.} \quad \text{.....(15)}$$

This should be equal to the sum $\sum_1^N \chi_r$ of the successive ionisation potentials of the atom. Estimates of the successive ionisation potentials of oxygen ($N = 8$), iron ($N = 26$) and silver ($N = 47$) have been tabulated by Hartree†. The totals, and the values of E given by (15) are shown in the following table, to which hydrogen ($N = 1$) and helium ($N = 2$) have also been added:

N	$\sum_{r=1}^N \chi_r$	$17N^{\frac{7}{3}}$	$\Sigma\chi_r/N^{\frac{7}{3}}$
1	13·5	17	13·5
2	80	86	15·9
8	2000	2200	15·6
26	33600	34000	16·8
47	138000	136000	17·3

The agreement is quite satisfactory. Since Thomas's theory holds only for atoms with a comparatively large number of electrons, we should expect $\Sigma\chi_r/N^{\frac{7}{3}}$ to tend to a limit with increasing N. This appears to be confirmed by the last column in the table.

This constant c_1, which we have determined as the difference between two large numbers, could be found more accurately by

* Corrected value, kindly supplied by Mr Thomas.

† *Proc. Camb. Phil. Soc.*, 22, 473, 1924.

Digitized by Google

integration if necessary. For the constant b_1 of (11) is proportional to v_0, i.e. to the self-potential of the distribution ρ at the origin, and so must be proportional to

$$\int_0^\infty \frac{\rho}{r} r^2 dr.$$

We readily find in fact that

$$\int_0^\infty \phi^{\frac{3}{2}} \xi d\xi = -\lim_{\xi \to 0} \left(\phi + \xi \frac{d\phi}{d\xi}\right) = -b_1.$$

6. *Summary.* Thomas's differential equation for the average field inside a heavy atom is analogous to Emden's equation for the polytropic equilibrium of a star. Emden's result that the total gravitational potential energy of a star is calculable once the differential equation has been solved is adapted to give the total electrostatic energy, and hence the total energy of binding, of an atom built on Thomas's model. This should be equal to the sum of the successive ionisation potentials. The total energy is found to be proportional to $N^{\frac{7}{3}}$, where N is the atomic number. The values found agree with Hartree's calculations of the successive ionisation potentials of certain atoms.

Digitized by Google

The Hall effect in Single Crystals of Iron. By Dr W. L. WEBSTER, Trinity College. (Communicated by Dr KAPITZA.)

[*Received* 20 May, *read* 25 July, 1927.]

Introduction.

In this paper the results of some measurements of the Hall effect in single crystals of iron are described. This work was undertaken to determine whether the Hall coefficient varied with the direction in the crystal for which it was measured. Such a variation had been observed in other magneto-electric phenomena—the change of resistance in longitudinal* and transverse† magnetic fields—and had thrown some light on the nature of these effects. It was thought possible, therefore, that the study of the Hall effect in single crystals might yield some fresh light on this unexplained phenomenon.

The results show that the Hall coefficient is independent of the direction in the crystal, and that the effect is similar to that observed in ordinary soft iron.

Measurements.

The Hall potential was measured as a function of the magnetic field in thin rectangular plates cut from single crystals. Copper leads were soldered to the ends and middle points of the sides of the plate. A steady current of one ampere was passed through longitudinally, and the Hall potential difference of the side electrodes due to the magnetic field measured by a potentiometer null method. As it was not possible to fasten the side electrodes exactly opposite each other, there was always a small potential difference due to the drop in potential along the plate. This was compensated by the potentiometer, the magnetic field was then applied and the change in the setting of the potentiometer necessary to restore balance was measured; finally the field was removed and the initial setting checked. This process was repeated several times. In this way values of the Hall potential were found for both directions of the field and of the primary current. The four figures obtained usually showed a systematic variation of about 5%; their average was taken as giving the correct value, the probable error being about 2%. (To record the balance, a galvanometer with a voltage sensitivity of 20 mms. per micro-volt was used.)

* *Proc. Roy. Soc.*, vol. 113, p. 196 (1926).
† *Proc. Roy. Soc.*, vol. 114, p. 611 (1927).

Digitized by Google

The variation of the four values of the Hall potential was probably due to the change of resistance of the iron caused by the magnetic field; as this potential change does not change sign with the magnetic field the error is eliminated by the averaging. No attempt was made to eliminate the Ettinghausen effect, as it did not seem possible that it could mask a variation of the Hall coefficient.

To prevent distortion of the plate by the magnetic field it was clamped between two plates of ebonite which were in turn wedged between the pole-faces of the magnet. The electro-magnet used had pole-faces 1·8 cms. in diameter and 0·6 cm. apart, and gave fields up to 22,800 gauss.

The plates used were cut out from the surface of some large crystals which were available. Plates of the required size about $\frac{1}{2}$ mm. thick were cut out and then ground down on gradually finer emery paper to a thickness of about 0·17 mm. Finally they were etched in nitric acid to remove any disturbed surface layers. The ratio of the length to the width of the plates was made greater than 3 : 2 to avoid any variation in the Hall coefficient* due merely to the use of plates that were too short in comparison with their lengths. The plates used were about 1·3 cms. long and 0·7 cm. in width.

Results.

The crystals from which the plates were cut were chosen to have their faces nearly parallel to a (100), (110), or (111) crystal plane, and the plates were cut out so that the directions of the primary current and the Hall potential should be as close as possible to the simple crystallographic axes of these planes.

Plate	Length (cms.)	Width (cms.)	Thickness (mm.)	Normal to Plate	Short Axis of Plate	Long Axis of Plate	Hall Coefficient
A	1·40	0·85	0·185	(100) − 6°	(100) − 6°	(100) − 6°	$+0{\cdot}97\times10^{-2}$
B	1·35	0·70	0·170	(110) − 8°	(110) − 7°	(100) − 1°	$+0{\cdot}98\times10^{-2}$
C	1·25	0·65	0·110	(110) − 8°	(100) − 1°	(110) − 7°	$(+1{\cdot}14\times10^{-2})$
D	1·30	0·85	0·165	(111) − 10°	(110) − 1°	(112) − 6°	$+0{\cdot}97\times10^{-2}$
Soft Iron							$+1{\cdot}13\times10^{-2}$

In the table are given the linear dimensions of the plates, and the crystal directions nearest the principal axes of the plates, the

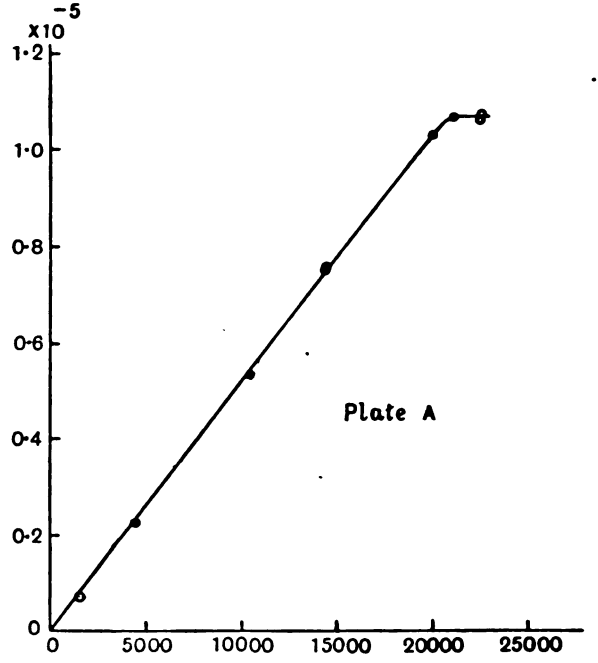
* Campbell, *Galvanomagnetic Phenomena*, p. 31.

Digitized by Google

deviation being shown. In the last column are the values of the Hall coefficient in absolute units calculated from the formula

$$R = \frac{Ed}{iH},$$

where E is the Hall potential, d the thickness of the plate, i the total current through the plate, and H the magnetic field. The probable error of these values is about 3%. The value of the coefficient for soft iron* is also given.



The curves showing the relation between the Hall potential and the magnetic field were much the same for all the plates, and only that for A is given. The potential increases linearly with the field up to about 21,000 gauss and then remains constant, the bend at the saturation point being sharper than in soft iron.

Discussion.

The results show that there can be no large variation of the Hall effect in single crystals. The values obtained for the different plates are constant within the experimental error; plate C, which

* Campbell, *loc. cit.*, p. 122.

Digitized by Google

gives a rather high value, was the first plate cut out and was badly strained in the process, so that this value cannot be trusted. It is interesting to note that the Hall coefficient given by this plate is in agreement with the value for soft iron.

The curves are very similar to those obtained for soft iron, the main difference being the sharper bend at the saturation point. This probably results from a more uniform intensity of magnetisation in the single crystals, the Hall potential being proportional to the intensity of magnetisation and not to the internal magnetic field.

No connection is apparent between this effect and the change of resistance of iron crystals in a magnetic field, so that it seems that the two phenomena must be quite different in origin.

My thanks are due to Prof. Sir Ernest Rutherford and Dr Kapitza for their interest in this work.

Digitized by Google

Excitation by High Velocity Electrons. By Mr E. G. Dymond, Fellow of St John's College.

[*Received* 20 May, *read* 25 July, 1927.]

Much attention has been focused on the excitation of atomic spectra by electrons whose velocities do not greatly exceed the critical value, and it is now possible to form a rough idea of the behaviour of the "excitation function," or probability that an electron of given velocity shall excite on impact, when the energy of the electron is only slightly greater than that required for excitation. But the form of this excitation function for much higher velocities is still in considerable doubt.

The most valuable work in this direction is that of Hughes and Lowe*, who investigated the variation in intensity of the various spectrum lines of helium, when bombarded by electrons of given velocities. Their method may however be objected to on two counts. It is necessary to assume that the light observed is caused only by impacts of electrons of the known impressed velocity, and that all secondary effects are excluded. The form of their apparatus was such that excitation by secondary causes, such as electrons reflected from the walls of the apparatus, was observed simultaneously with the primary effect.

Secondly, it is necessary that not only the current should remain constant throughout the course of the experiment, but that the current density also should not vary in the space in which observations are made. Within a certain region of velocities this last condition appears impossible to fulfil, at least without extreme additional complications, and it is unfortunate that this region is precisely that in which Hughes and Lowe worked.

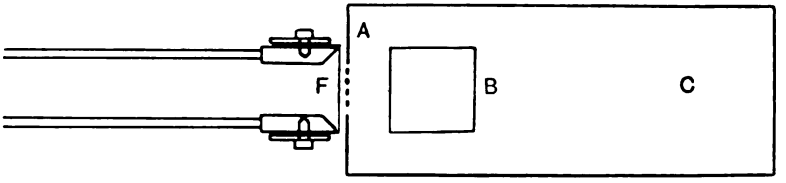


Fig. 1.

Accordingly it was decided to repeat these experiments with a form of apparatus as little open to objection from these quarters as possible. The tungsten filament as source of electrons was replaced by an oxide-coated platinum strip, *F* (Fig. 1), about 10 mm. × 10 mm.

* *Proc. Roy. Soc.* 104, p. 480 (1923).

Digitized by Google

in size. To reduce the current necessary to heat it and to localise the heating effect two fine V-shaped slits were cut in the centre, reducing the effective width at this point to 5 mm. Immediately opposite the strip and 1 mm. from it was placed a nickel disc, *A*, pierced with a 5 mm. hole covered with fine nickel gauze. The disc formed one end of a nickel cylinder, *C*, 5 cm. long and 2 cm. in diameter. A hole, *B*, in the wall allowed of observation of the interior.

With this type of source it is found that when a potential is applied between *F* and *A* the electrons proceed outward normally from the strip, and owing to the localisation of the heating, form a sharply defined beam proceeding downwards into the cylinder. This effect is no doubt partly due to the uniform direction of the field imposed by the geometrical arrangement, but also, in the presence of a gas, to the positive space charge "sheath" formed by the positive ions, in which all the potential drop occurs. Where a gas is present at low pressure this beam is visible as a luminous streak with sharp edges, surrounded by a weak diffuse glow which seems to fill the whole cylinder. This glow is to be ascribed to excitation by secondary causes; but if the slit of the observing spectrograph is placed so that it is perpendicular to the direction of the beam, a line is obtained on the plate which is strong in the centre and weak at the edges. By taking the difference in intensity between the outer and inner portions, the effect of the secondary excitation can be eliminated.

The form of the beam undergoes changes with varying impressed velocities. It is at first very diffuse. In helium with 50 volts between anode and cathode, it forms a cone of not very wide angle, which narrows to a parallel beam at 100 volts, and maintains this form unaltered in dimensions to 1800 volts, the highest potential applied. It is this alteration in the form of the beam, and consequent variation in the current density, which makes measurements of intensity below 100 volts unreliable. It should be possible by integrating the intensity over the width of the beam to carry the measurements to considerably lower velocities, but this so far has not been attempted. The constant form of the beam above this limit, however, allows one to hope that the results here may have some real meaning.

A difficulty inseparable from the use of an oxide-coated source, especially at high voltages, is the inconstancy of the emission. The bombardment of the source by high velocity positive ions has a sensitising effect, and the emission, if not kept constant by reducing the heating current, increases enormously. The magnitude of this effect may be gauged from the fact that at the start of a run at, say, 500 volts, the filament was glowing bright red. In the course of 2 hours it was necessary to reduce the heating

Digitized by Google

current to such an extent that the source was at the end only just visible to dark-adapted eyes. The variation in the emission current was kept within 2 per cent. by control of the heating current, a variation below the errors arising from the photometry of the plates.

The gas used was helium and was purified in the usual way over heated copper oxide and charcoal cooled in liquid air; all the metal parts were glowed to a red heat in vacuum before assembly, and the whole apparatus, with the exclusion of a ground glass joint which supported the electron source, was thoroughly baked to 450° C. before a run was made. A wide tube filled with charcoal cooled in liquid air was in immediate communication with the apparatus and served to remove traces of foreign gases evolved during the course of a run. That these precautions for assuring pure helium were adequate may be seen from the fact that no trace of any lines foreign to helium was ever found, even with the longest exposures.

The gas pressure varied in different exposures from 0·09 to 0·05 mm., and the total current between anode and cathode was 900 microamperes.

A Hilger constant deviation spectrograph with glass prism was used for observing the radiation, which was consequently limited to the visible region. In order to facilitate the measurement of the blackening of the plates, after each exposure at a definite electron velocity, the drum rotating the prism was turned by a small amount, so that several images of the same line appeared side by side, each due to excitation at a different velocity. It is fortunate that the spectrum of helium allows this to be done without overlapping of the images of different lines. By this means all the data for a single line are collected together in a small area of the plate and may be obtained by one setting of the photometer.

There was photographed on the same plate the spectrum of helium from a Geissler tube excited by a transformer from an alternating supply which was maintained constant during the exposure. Images of varying intensity were recorded in the same way as described above, the intensity being controlled by means of wire screens, kept in continuous motion by pendulums, interposed in the path of the light. The absorbing power of the screens was measured directly on a photometer.

In all, three photometers were used in the research. The first was a photographically recording thermocouple instrument at Bryn Mawr College, Pennsylvania, constructed by Prof. James Barnes, the second a non-recording thermocouple instrument at the Palmer Laboratory, Princeton, New Jersey, and the third a null-method photoelectric cell instrument of the Dobson type, in which the blackening is directly compared with a wedge, at the

Digitized by Google

Cavendish Laboratory, Cambridge*. The plates used were Ilford special rapid panchromatics and the time of exposure was usually $1\frac{1}{2}$ hours.

Discussion of Results.

The variation in intensity of the various visible lines of the helium spectrum is shown in Fig. 2. Each line is plotted against

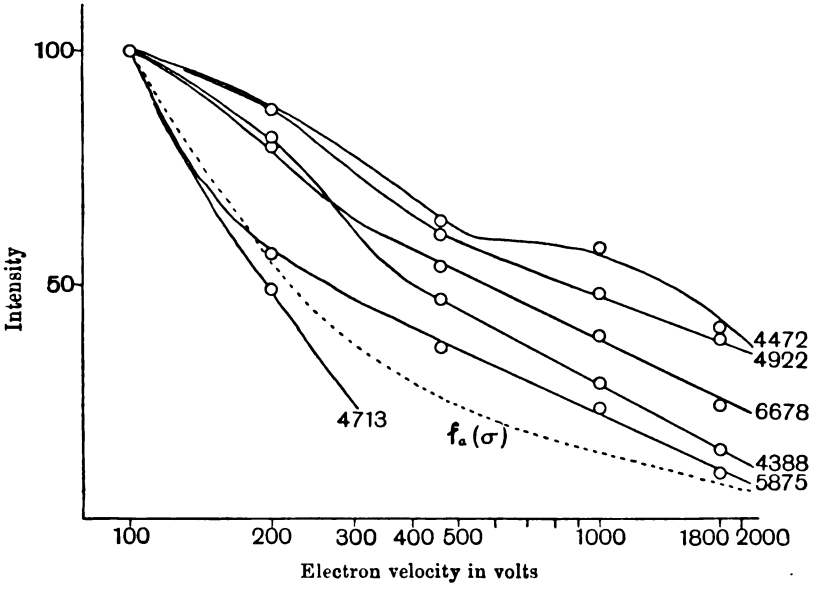


Fig. 2.

a standard intensity of 100 at 100 volts exciting velocity. Except for qualitative purposes comparison between the intensities of the various lines is useless unless the variation of sensitivity of the plate with wave length is known. The voltage of the exciting electron is plotted on a logarithmic scale for the sake of convenience. The curve marked $f_a(\sigma)$ is the theoretical curve for the efficiency of excitation of atomic hydrogen† derived by Born‡.

* I am deeply indebted to Prof. Barnes for his courtesy in allowing me to use his instrument and also to Mr C. Thomas for his help in measuring some of the plates on the Princeton photometer.

† In conversation with Prof. Born I have gathered that he expects no great change in the form of the function when applied to atoms other than hydrogen.

‡ Born, "Zur Wellenmechanik der Stossvorgänge," *Göttinger Nachrichten*, p. 146 (1926).

Digitized by Google

The following lines have been investigated:

Wave Length	Series Notation
6678 Å.	$2\,^1P—3\,^1D$
4922	$2\,^1P—4\,^1D$
4388	$2\,^1P—5\,^1D$
5875	$2\,^3P—3\,^3D$
4713	$2\,^3P—4\,^3S$
4472	$2\,^3P—4\,^3D$

No systematic difference in behaviour between the lines of the single and triplet systems is to be noticed, which does not confirm the observations of Hughes and Lowe, in the region of lower velocities, that the lines of the triplet (ortho-helium) system fall off in intensity more rapidly. The line 4713 is to be excepted. This line reduces in intensity with far greater rapidity, and at the higher voltages is not found at all. Further it is to be noticed that over a large range the curves for the various lines are linear and run parallel, indicating that their excitation functions are the same except for a numerical factor affecting the velocity. The theoretical curve of Born gives the general tendency of the function, but cannot be said to give a very satisfactory fit.

The similarity in behaviour of the two systems of lines is surprising when one considers that for the excitation of the ortho-helium spectrum the atom has to undergo a transition forbidden most stringently alike by the old correspondence principle and by the new view of Heisenberg*. An inspection of the plate shows further no very marked difference in intensity between the lines of the two series. At 1800 volts, the relative blackening of $\lambda 6678$ and $\lambda 5875$, two lines of the same notation in the two series, is 2·5 : 1. The relative sensitivity of the plate to these two wave lengths is unknown but should favour the yellow rather than the red line. For another pair, $\lambda 4922$ and $\lambda 4472$, the blackening is nearly the same.

Now it is possible to obtain information as to the relative efficiencies of excitation of the two series from a completely different angle. By a development of the electron collision method the author† has been able to study the loss of energy of electrons of predetermined velocity on making a single collision with a helium atom.

Each energy state of the atom gives rise to a sharp peak in the curve (Fig. 2 of "Electron Scattering"), the height of that peak giving the number of electrons which in collision have lost energy of that particular amount. Now for an initial velocity of 49 volts

* *Zeit. f. Phys.* 28, p. 411 (1926).

† Dymond, *Physical Review*, 29, p. 433 (1927); this will be referred to as "Electron Scattering."

Digitized by Google

the principal energy loss results in excitation to the $2\,^1S$ state of the singlet system, while excitation to the triplet state, represented by $2\,^3S$, is only just detectable. For higher velocities, and the investigation was carried up to 600 volts, no trace of excitation to the $2\,^3S$ state is detectable. The resolution of the apparatus was not such as to distinguish higher triplet from singlet states, but it is highly unlikely that the higher members of the series should be excited so much more strongly than the first member. It should be remembered that for energies close to the critical potential it is the $2\,^1S$ state which is most excited. This behaviour of the two types of excitation is in fact explicable on the new resonance theory of Heisenberg*. On this view, when the two electrons composing the helium atom are perfectly equivalent, no transition whatever between the two series of states is possible, and such transitions are only possible when the equality of the two electrons is disturbed. Now Heisenberg assumes that the axes of the electrons shall either be parallel or antiparallel. The approach of an outside electron introduces an interaction between it and the electrons of the atom, which will in general render their axes no longer parallel. The equivalence of the two electrons is now disturbed and a transition from the singlet to the triplet system of states can take place. But the faster the colliding electron the less time do the perturbing forces have to play, and the smaller will be the distortion of the axes, and the smaller the probability of transition from one system to the other.

We may consequently draw the conclusion from the experimental evidence presented in "Electron Scattering" that the excitation of the triplet system of states decreases very rapidly with increasing velocity of the colliding electrons, with respect to that of the singlet system, and in fact for velocities of 100 volts or over is entirely negligible. And yet it is found that the light emitted, due to transition within the triplet system, is of the same order of intensity as that of the singlet system, and that the two systems of lines behave in approximately the same way.

It is possible to explain this paradox however, if we assume that the radiation emitted when an atom is struck by a fast electron, is not due to quantised excitation in the ordinary sense of the word, but is of the nature of a radiation of recombination, during the act of ionisation†. According to this view, there is no

* *Loc. cit.*

† I am deeply indebted to Professor Franck for pointing out to me the possibility of this type of excitation of light. In the course of correspondence he has very kindly informed me of the possibility of recombination of a positive ion with the electron which has been removed to form it, before that electron has had time to recede from the sphere of influence of the ion. This of course is a radically different process from the recombination of an ion with a "free" electron, for which no sound experimental evidence has been adduced.

Digitized by Google

a priori reason why the electron on recombining should fall rather into the par- than into the ortho-helium configuration, and we should expect roughly equal intensities for the two series, as is found. It is not necessary to assume that this type of excitation is favoured by high velocity of the original electron, but it appears to become noticeable rather by the cessation of the normal quantised process, which at lower velocities contributes the bulk of the radiation.

Reference to the results shown in "Electron Scattering" raises however a doubt as to this interpretation. It is shown there that for higher initial velocities a group of electrons appears which has lost on collision an indeterminate amount of energy, exceeding that required for excitation. For an initial velocity of 400 volts, this group seems to start at an energy loss of 23 volts and is appreciable to some 200 volts loss. It is possible, though unlikely, that the resolving power of the apparatus was insufficient to distinguish a 23 volts loss from one of 25 volts, that required for ionisation, and that this "continuum" is to be ascribed to ionisation. Other points are raised in "Electron Scattering" which do not favour this view. Further experiments are under way, which it is hoped will decide the exact point where this continuum begins.

If we adhere to the view that continuous loss of energy begins at a point below the ionisation potential, we are forced to conclude that when the time of collision is very short compared with the natural frequencies associated with the motions in the normal helium atom, as is the case here, the atom may not assume immediately one of the normal series of states, but may do so, cascade fashion, from one of a continuous series. The intensity relations of the two normal series when excited in this manner are naturally obscure, but it seems conceivable that a way to explain the anomalies at higher voltages of excitation may be found in this direction, if the results on the energy losses given in "Electron Scattering" are upheld.

If we accept either of these explanations of the behaviour of the intensities it is not surprising that the theoretical curves of Born do not fit more closely the experimental results, as he considers only the case of direct excitation to the upper state. It should be noted that his curve lies below those of all the lines, which is in agreement with the view that another process is also taking place.

I am much indebted to Prof. Karl T. Compton for extending to me the privilege of working in the Palmer Laboratory of Princeton University, where this work was carried out, and also to the International Education Board for their support by a Fellowship.

Digitized by Google

A Valve Amplifier for Ionisation Currents. By C. E. WYNN-WILLIAMS, Trinity College. (Communicated by Prof. Sir E. RUTHERFORD.)

[*Received* 20 May, *read* 25 July, 1927.]

Attempts have often been made, by means of valve amplifiers, to magnify small ionisation currents so that they can be measured by galvanometers of comparatively low sensitivity. The instabilities in the amplifying system, generally due to fluctuations in the voltage of the batteries supplying the valve currents, cause erratic changes in the position of the zero of the galvanometer, rendering accurate working impossible. Consequently such methods have not come into general use. But yet the advantages arising from the substitution of a low sensitivity, short period galvanometer for a high sensitivity electrometer having a long time of swing, or a high sensitivity galvanometer, are so self-evident that any method of making a valve amplifier more reliable from the point of view of a steady zero is of interest.

In the present paper, a system of "compensation," involving the use of only two valves, is described, by which the trouble of an unsteady zero is overcome to such an extent that it is possible to measure ionisation currents of the order of 10^{-12} to 10^{-13} amperes by means of a galvanometer of low sensitivity. If desired, a table or recording instrument may be used for special work. The current amplification factor for the apparatus is of the order of 10^5.

While the advantages of the system are discussed later, it may here be said that besides enabling a high sensitivity instrument to be replaced by a more robust and easily adjusted low sensitivity galvanometer, the compactness of the apparatus enables it to be moved from place to place, and quickly set up when required. It is thus specially suitable for lecture demonstrations, it being easily possible to illustrate with it such phenomena as the relative degrees of ionisation produced by α, β and γ rays, stopping powers, Bragg curves, decay curves, etc. The principle of compensation can also be applied to other valve apparatus where great steadiness of zero is required.

In addition to describing the actual methods of obtaining a steady zero, the characteristics of the apparatus are included in the account as a guide to any who desire to use this method for measuring ionisation currents.

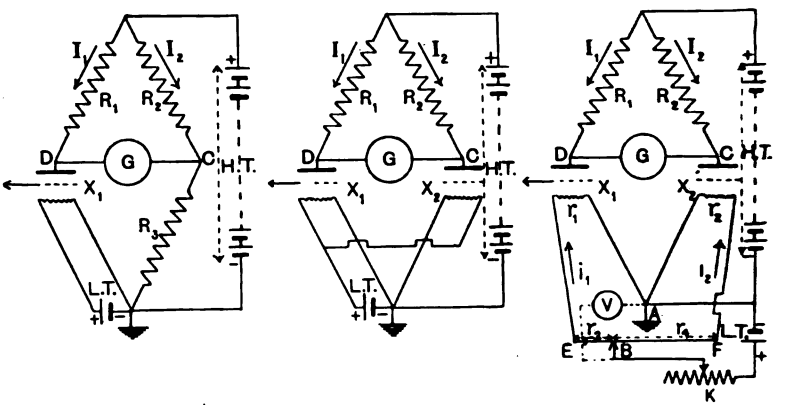
Old Methods.

Valve methods of measuring ionisation currents generally depend upon the relaying action of the thermionic valve. The small current to be measured is sent through a high resistance,

Digitized by Google

and the difference of potential set up across the latter is applied to the grid and earth terminals of a voltage amplifier. Alternatively, the ionisation current may be driven on to the insulated grid of a three-electrode valve, gradually raising its potential (in virtue of its small capacity) until the ionisation current is in equilibrium with the leakage current from the grid. In either case, the final result is the same—the change in grid potential causes the plate current in the valve to change, the increase or decrease (read by means of a galvanometer) being used to measure the ionisation current. A change of one volt in the grid potential usually results in a change in the plate current of the order of a milli-ampere.

The alternative method (in which the insulated grid is used) is the more sensitive, although this is only a particular case of the other method, for a certain amount of "grid-leakage" is always present, and may be regarded as the high resistance through which the ionisation current is being driven when a state of equilibrium has been established between the ionisation current and the leakage current.



(*a*) (uncompensated) (*b*) (uncompensated) (*c*) (compensated)

In each case, the grid of valve *D* is connected to the ionisation chamber, while that of *C* is free.

Fig. 1.

The usual circuit employed is shown in Fig. 1(*a*). In series with the plate *D* of a valve is placed a resistance R_1, of the order of 10,000 ohms, while two other variable resistances, R_2 and R_3, are connected in series across the high tension battery H.T., supplying the plate current. The ionisation chamber is connected to the grid, and the negative pole of the low tension battery L.T., supplying the filament heating current, is usually connected to earth. The galvanometer, *G*, is connected between the anode and the junction of R_2 and R_3.

Digitized by Google

Denoting the currents in R_1 and R_2 by I_1 and I_2, the galvanometer will be undeflected if the relation $R_1 I_1 = R_2 I_2$ is satisfied. Alternatively, denoting the impedance of the valve to the plate current by X_1, and regarding the system as a Wheatstone bridge, balance will be obtained when $R_1/R_2 = X_1/R_3 = I_2/I_1$. Any change in the potential of the grid (produced by the ionisation current) alters the value of I_1 (or of X_1) and so throws the system out of balance, giving rise to a galvanometer deflection which, it can be shown, is proportional to the change of grid potential*.

While such a system may have a large current amplification factor, an early limit is set to the minimum current that can be measured by it *by the unsteadiness of the zero*, and no advantage is to be gained by using an extremely sensitive galvanometer unless the system can be stabilised.

The unsteadiness of the zero is mainly due to the small changes of the plate current I_1 produced by small fluctuations of the voltage of the L.T. or H.T. batteries. In other words, the impedance X_1 of the valve is not constant, but a function of the L.T. and H.T. battery voltages. Hence, if R_1, R_2 and R_3 [Fig. 1(*a*)] are constant, any change in X_1, produced by battery voltage fluctuations, will result in the system becoming unbalanced.

This unsteadiness can be overcome to a certain extent by using another valve [Fig. 1(*b*)] similar to the first, and of impedance X_2, to replace R_3 [of Fig. 1(*a*)]. The first valve only is used as an amplifier, the second being idle. With two perfectly matched valves, the debalancing effects produced by battery voltage fluctuations would cancel out. No two valves, however, have characteristics sufficiently similar to give perfect compensation in this manner, although this circuit is an improvement on the one shown in Fig. 1(*a*).

Compensated System.

The foregoing will have made it clear why valve methods of measurement, while they have been regarded with great interest as laboratory experiments, have not seriously been considered as possible rivals to the quadrant electrometer or high sensitivity galvanometer.

By means of the method described in subsequent pages, a valve bridge similar to the type shown in Fig. 1(*b*) can be changed from what might perhaps be regarded as a wireless toy, whose action was always accompanied by a considerable degree of uncertainty, to a physical measuring instrument, capable of giving useful and consistent results for a variety of purposes where an electrometer would normally be used.

* Provided that the valve is operated on a suitable portion of its triode characteristic.

Digitized by Google

The necessary modification seems at first sight to be so superficial as to render such a claim absurd, it consisting of merely 30 cm. of Eureka wire, *EF*, connected to the filaments of the valves [see Fig. 1(*c*)]. But a consideration of the theoretical side of the question shows that this is not the case, the inclusion of the slide wire *EF* being the most natural and simplest way of enabling certain theoretical conditions to be realised experimentally, and not a haphazard addition to an existing older circuit.

It may be wondered why, if the circuit is so simple, it has not been already tried in the past. But again, even if it had been tried, the full advantages of it are only realised when it is properly adjusted in accordance with theoretical considerations. In other circumstances, little or no advantage will be gained over the older circuit of Fig. 1 (*b*).

Fortunately, these adjustments, while very necessary, are extremely simple to carry out, and when once made, do not have to be altered except on rare occasions. As in the older circuit of Fig. 1 (*b*), one valve only is used to amplify, the other being employed to make the system symmetrical. Any slight differences in the characteristics of the two valves are compensated for over a small, though sufficient range of working voltages, so enabling the advantages of a pair of perfectly matched valves to be realised experimentally.

As so much depends upon the theoretical considerations, a full discussion is given in the next sections. Meanwhile, in passing, attention is directed to the curves of Figs. 4(*b*), 4(*c*) and 4(*d*). These are records showing the degree of steadiness of zero attained for three systems; Fig. 4(*b*), a *compensated* system as described here [Fig. 1(*c*)], and Figs. 4(*c*) and 4(*d*), *uncompensated* systems similar to those shown in Figs. 1(*a*) and 1(*b*). The advantage of this method of using valves is well brought out by a comparison of the curves.

Compensation for Low-tension Battery Voltage Fluctuations.

Assuming that the H.T. voltage is constant, the effects of small changes produced in the filament currents by a change in the low tension battery voltage is to alter the electronic emission of the filaments, and thus give rise to changes in the plate currents (I_1 and I_2). If the filament circuit can be so arranged that while I_1 and I_2 vary, their ratio is always kept constant during these changes, the galvanometer deflection will remain undisturbed by small L.T. battery voltage fluctuations.

For I_1/I_2 to be constant, we have

$$\frac{1}{I_1}\frac{dI_1}{dv} = \frac{1}{I_2}\frac{dI_2}{dv},$$

Digitized by Google

where v is the L.T. battery voltage. That is, the *percentage changes* in the plate currents must be equal. The two filament currents being denoted by i_1 and i_2, the above condition may be written:

$$\frac{1}{I_1}\frac{dI_1}{di_1}\frac{di_1}{dv} = \frac{1}{I_2}\frac{dI_2}{di_2}\frac{di_2}{dv}.$$

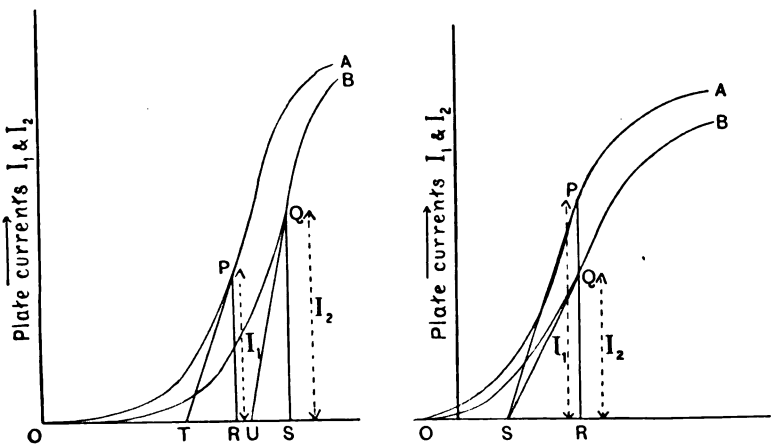
While the first two factors of each side of the equation

$$(dI_1/I_1di_1 \text{ or } dI_2/I_2di_2)$$

depend upon the characteristics of the valves, and the experimental conditions, the last (di_1/dv or di_2/dv) can be varied at will by a suitable adjustment of resistance in series* with the filaments (i.e. the slide wire EF).

For any given experimental conditions, dI_1/I_1di_1 and dI_2/I_2di_2 will have definite values. If, therefore, the values of di_1/dv and di_2/dv can be adjusted by means of the slide wire, until the ratio $\frac{di_1}{dv}\Big/\frac{di_2}{dv}$ is equal to the ratio $\frac{1}{I_2}\frac{dI_2}{di_2}\Big/\frac{1}{I_1}\frac{dI_1}{di_1}$, compensation will be effected, and L.T. battery voltage fluctuations will not upset the balance.

That this adjustment of the ratio of $\frac{di_1}{dv}\Big/\frac{di_2}{dv}$ is possible may be



Filament currents i_1 and i_2.
(*a*)

H.T. voltage V.
(*b*)

Note. These curves are illustrative only, and do not refer to specific conditions.

Fig. 2.

* In this paper the filaments are assumed to be connected in parallel. By using *shunt* resistances, compensation can also be effected when the filaments are connected in series.

Digitized by Google

seen by considering the form of the emission curves for the valves. Fig. 2(*a*) shows the general form of the curves for two valves of the type employed in this apparatus, the plate current being plotted as a function of the filament current, and the H.T. voltage being assumed constant.

The shapes of the curves depend upon the type of valve employed, and the value of the H.T. voltage, etc. If two precisely similar valves were employed, the curves AO and BO would be coincident. Actually, they lie very close together, but are shown well separated in the figure to avoid confusion.

Let P and Q be two points on the curves, corresponding to the values of the filament current and plate current of the two valves under a given set of conditions. The perpendiculars PR, QS represent I_1 and I_2, while the slopes of the tangents PT, QU to the curves at P and Q will be measures of dI_1/di_1 and dI_2/di_2 (i.e. $dI_1/di_1 = PR/TR$ and $dI_2/di_2 = QS/US$). Hence

$$\frac{1}{I_1}\frac{dI_1}{di_1} = \frac{1}{PR}\frac{PR}{TR} = \frac{1}{TR}$$

and

$$\frac{1}{I_2}\frac{dI_2}{di_2} = \frac{1}{QS}\frac{QS}{US} = \frac{1}{US},$$

so that the ratio $\frac{1}{I_2}\frac{dI_2}{di_2}\Big/\frac{1}{I_1}\frac{dI_1}{di_1}$ (to which $\frac{di_1}{dv}\Big/\frac{di_2}{dv}$ must be made equal for compensation to be effected) is represented in the figure by the ratio TR/US.

This fraction, TR/US, can obviously have values ranging from zero to infinity, depending upon the positions of P and Q. In practice, it is endeavoured to operate the valves at points situated on fairly straight portions of their respective emission curves, as the compensation then holds over a greater range of L.T. voltage fluctuation. For a pair of similar valves the straight portions of the curves will lie close together, so that P and Q occupy approximately similar positions on their respective curves. Under these conditions, the value of the ratio TR/US is usually very close to unity.

The adjustment of $\frac{di_1}{dv}\Big/\frac{di_2}{dv}$ to this value is effected—as described in the next section—by altering the position of the slider B on the slide wire EF [see Fig. 1(*c*)]. Assuming the resistances of the filaments to be constant, and equal to r_1 and r_2, and of the portions of the slide wire in series with each to be respectively r_3 and r_4, it is clear that

$$i_1 = \frac{v}{r_1 + r_3}, \qquad i_2 = \frac{v}{r_2 + r_4}$$

Digitized by Google

(v being the L.T. battery voltage), so that

$$\frac{di_1}{dv}\Big/\frac{di_2}{dv} = \frac{r_2+r_4}{r_1+r_3} = \frac{r_2}{r_1}\left(\frac{1+\dfrac{r_4}{r_2}}{1+\dfrac{r_3}{r_1}}\right).$$

Thus, by altering the position of the slider on the wire, the ratio of r_3 to r_4 may be varied until $\frac{di_1}{dv}\Big/\frac{di_2}{dv}$ is of the value required for compensation to be effected.

Practical Method of Compensating.

In effecting compensation it is not necessary to draw the characteristic curves for the valves and to determine from them the most suitable position of the slider B on the wire, and the value of the filament voltage, etc. A simple method of trial and error suffices, and is more convenient. In the first experiments, however, in order to investigate the matter properly, the characteristic curves were obtained.

Referring to the diagram of Fig. 1(c), the procedure of compensating is as follows*. The slider B (which divides the wire into the two resistances r_3 and r_4) is placed at the centre of the wire, and the voltage across AB is raised, by cutting out some of the resistance K, until it is near the normal working voltage of the valves (i.e. about 1·8 volts for dull-emitter valves). R_1 and R_2 are then adjusted until the system is balanced (i.e. until $R_1/R_2 = I_2/I_1$), as shown by zero deflection of the galvanometer. (For this part of the work a table galvanometer, of sensitivity about 1 or 2 divisions per micro-ampere, is preferable to a reflecting galvanometer.) A small definite change of a few per cent. is then made in the voltage across AB, and the amount of debalancing, i.e. the number of divisions of the galvanometer deflection, observed. B is then moved to one end of the wire, and the operation of balancing and observing the galvanometer deflection for a given change in the

* While in this method, the value of $\frac{di_1}{dv}\Big/\frac{di_2}{dv}$ is adjusted to equal the existing value of $\frac{1}{I_2}\frac{dI_2}{di_2}\Big/\frac{1}{I_1}\frac{dI_1}{di_1}$, the converse method is also possible, and sometimes more convenient. The ratio of $\frac{di_1}{dv}\Big/\frac{di_2}{dv}$ is fixed by clamping down the slider B in one position, and the characteristic curves are searched by raising the voltage across AB until a pair of points P and Q are found for which the ratio $\frac{1}{I_2}\frac{dI_2}{di_2}\Big/\frac{1}{I_1}\frac{dI_1}{di_1}$ is equal to the predetermined ratio $\frac{di_1}{dv}\Big/\frac{di_2}{dv}$, indicated by a reversal in the galvanometer deflection.

Digitized by Google

voltage across AB is repeated. The same procedure is followed with B at the other end of the wire.

It may be found that at one end of the wire the galvanometer deflection for a given change in voltage across AB is *increased*, while at the other end it is either *decreased* or *reversed*. Should a reversal be obtained at one end, a point may be found (by trial and error) between that end and the centre of the wire, at which no debalancing occurs for a small change in the voltage across AB (i.e. at which the system is compensated). Should a reversal not be obtained at either end of the wire, the operation should be repeated with values of the filament voltage above or below the normal one.

Degree of Compensation Obtainable.

The degree of compensation obtainable, that is, the range of L.T. voltage fluctuation which produces no appreciable debalancing, depends on what portions of the emission curves P and Q lie upon. It is clear that if they lie on comparatively straight portions of the curves, for which dI_1/di_1 and dI_2/di_2 are constant, a larger L.T. voltage fluctuation can be compensated for than otherwise. By a suitable manipulation of the slider B and the rheostat K, it is possible to move P and Q along their respective curves, while still keeping the system compensated, until the optimum position is found, the relative degree of compensation in each case being measured by the ratio of the change in voltage applied to AB, to the amount of debalancing it produces.

To convey some idea of the results obtainable, it may be said that with the valves used in this apparatus, no difficulty should be experienced in compensating the system so that a change of about 0·05 volt across AB, in a mean value of between 1·60 and 1·90 volts, produces a change of less than 0·25 micro-ampere in the galvanometer current. This means that the ratio of I_1/I_2 is kept constant to about one part in ten thousand for a 2·5 per cent. change in the voltage of the low tension battery.

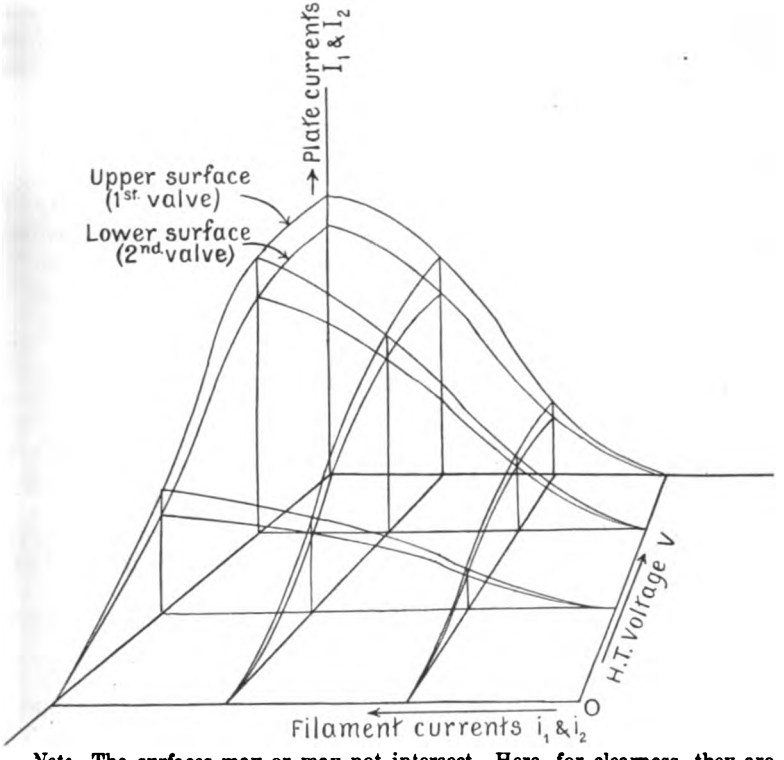
Compensation for H.T. Battery Voltage Fluctuations.

An investigation of the conditions for compensation for small fluctuations of the H.T. battery voltage results in a similar expression to the one relating to the L.T. battery voltage, i.e. $\frac{1}{I_1}\frac{dI_1}{dv} = \frac{1}{I_2}\frac{dI_2}{dv}$, v being now the H.T. voltage, the L.T. voltage being assumed constant.

On plotting the relation between I_1 or I_2 and v under normal working conditions, curves of the form shown in Fig. 2(*b*),

Digitized by Google

similar to the well-known diode emission curves, are obtained. Since both valves are operated at the same H.T. voltage, the two points P and Q on the curves, at which the valves are operated, must lie on the same perpendicular PQR. Subject to this condition, it may be shown that the relation $\frac{1}{I_1}\frac{dI_1}{dv} = \frac{1}{I_2}\frac{dI_2}{dv}$ only holds if the two tangents PS, QS, to the curves at the points P and Q



Note. The surfaces may or may not intersect. Here, for clearness, they are shown separate. Fig. 2 (*a*) represents a section perpendicular to the axis of V, and Fig. 2 (*b*) one perpendicular to the axis of i_1 and i_2.

Fig. 3.

intersect at some point S on the axis of v. Whether such a pair of points P and Q, having a common point of intersection of their tangents on the axis of v, exists depends upon the shapes of the curves. If it does, H.T. battery voltage compensation is theoretically possible, although the value of v in such a case may not be suitable for practical purposes. The shape of the curves can be modified to some extent by altering the values of R_1 and R_2 while still

Digitized by Google

keeping their ratio R_1/R_2 constant, which may or may not so alter the positions of P and Q as to make the corresponding value of r a suitable one for practical purposes.

Since the plate currents I_1 and I_2 are functions of both the H.T. and L.T. battery voltages, the curves shown in Figs. 2(a) and 2(b) must be regarded as perpendicular sections of two surfaces in three dimensions—one surface relating to each valve—similar to those shown in Fig. 3. While therefore individual compensation for either L.T. or H.T. voltage fluctuations may be separately effected, simultaneous compensation * for both these effects, which would mean exploring the surfaces of Fig. 3 for suitable working points, would be a very difficult and tedious process.

In practice, the debalancing produced by H.T. battery voltage fluctuation is not so serious as that produced by fluctuations in the L.T. voltage, and may generally be ignored if a rather larger capacity accumulator than is usual is employed for the H.T. battery. Under such conditions, low tension battery voltage compensation alone renders the zero sufficiently steady for currents of the order of 10^{-12} or 10^{-13} amperes to be dealt with.

Choice of Valves.

The first consideration in the choice of valves is the degree of amplification obtainable. It is assumed for simplicity that the total capacity and insulation of the ionisation chamber and grid is approximately the same whatever type of valve is employed, so that a given ionisation current causes the same change of grid potential in various cases. It may be shown that a change of grid potential E produces a change of galvanometer current equal to

$$E\frac{\mu}{2X+G\left(1+\frac{X}{R}\right)},$$

μ, X, R, and G being respectively the voltage amplification factor and impedance of the valves, the anode resistances, and the galvanometer resistance. (For simplicity, the system is assumed symmetrical.) If, as is usually the case, G is small compared with

* Where great steadiness of zero is required, and a large capacity H.T. battery is available, *simultaneous* compensation may be effected by using a common L.T. and H.T. battery. The slider B [Fig. 1(c)] should be connected through a resistance of about 150 to 300 ohms to the positive pole of the H.T. battery. As before, compensation is effected by varying the position of the slider B on EF until a small change in the voltage of the battery produces no debalancing. In this case, the characteristic curves on which P and Q lie are represented by a diagonal section of the surfaces of Fig. 3, the L.T. and H.T. being both represented proportionally on the horizontal axis. With this system, the size of the common 60 volt battery required to supply a filament current about a hundred times the value of the plate current is rather a drawback.

Digitized by Google

X, this may be written $E\mu/2X$. The greatest current magnification will therefore be given by a valve having a large value of μ/X, or mutual conductance: a power valve.

Another consideration is the type of filament employed. Dull emitter valves, operating at comparatively low filament temperatures, are obviously more suitable than the old bright emitter valves, not only on account of the smaller heating current required, but also owing to the fact that the smaller changes in the temperature produced by changes of filament current give rise to much smaller changes in the resistance of the filament, thus approximating more nearly to the assumption made on p. 816 that the resistance was constant.

Valves which give good results are the "Marconi 215" and "Osram 215," which normally require a filament voltage of 1·8, and a filament current of 0·15 ampere per valve. A two-volt accumulator of at least 30 ampere-hours capacity should be used for the L.T. battery. For the H.T. battery, a 60 volt 2·5 ampere-hour accumulator, preferably of the "block" type, is suitable. While the H.T. voltage may be increased to 120 with these valves, no advantage is gained unless large changes of grid potential, of the order of volts, have to be dealt with.

The average values of μ and X for the valves are 6·25 and 6·250 ohms. When a galvanometer of 50 ohms resistance, and 10,000 ohms anode resistances were employed, a milli-volt change in the grid potential gave rise to a galvanometer current of 1/3·15 micro-ampere, a good agreement with the value of approximately $\frac{1}{2}$ micro-ampere as calculated from the makers' average data.

Arrangement of Apparatus.

It is found convenient to have as much as possible of the apparatus mounted on a baseboard. Compensation then being once effected, the apparatus will remain in that state, and can be conveniently carried from place to place and quickly connected up as required.

The two valves fit into holders on an ebonite panel in an earthed metal box, screened leads to the grids of either valve being brought into the box by metal tubes. Connections to the anode resistances, batteries, etc. which are situated outside the box, are made with ordinary insulated wire. While such a high degree of insulation as is usual with electrometers is not required in this case, care should be taken that the insulation of the grid is not unduly lowered by poor quality ebonite, etc.

The slide wire EF consists of about 30 cms. of Eureka wire No. 28, having a resistance of between 1 and 1·5 ohms, and the sliding contact B is so arranged that it can be secured to the base-

Digitized by Google

board with wax when compensation has been effected. A scale under the wire is a convenience.

The anode resistances consist of two resistance boxes, each of 10,000 ohms, and a 500 ohm variable rheostat. The sliding contact of the latter serves as the common junction of the anode resistances, connected to the positive pole of the H.T. battery, while the two ends of the resistance are connected, through the 10,000 ohm boxes, to the anode terminals of the valves. By this means, the boxes being adjusted to the nearest 500 ohms, the system may be rapidly balanced by moving the sliding contact on the variable rheostat. This gives a convenient method of controlling the position of the zero of the galvanometer.

Most of the loose components (i.e. resistance boxes, voltmeters, etc.) can be conveniently situated on the baseboard and connected to appropriate terminals on the latter. The earth terminal should be connected to the point A [Fig. 1(c)] and also to the metal box surrounding the valves.

Precautions to be observed.

The zero of a new pair of valves, even after compensation has been effected, is very unsteady for some time. This is due to the fact that, the filaments being "thoriated," a state of equilibrium has to be established at their surfaces. After the system has been in operation for about 50 hours under normal conditions, the zero will be found to be very much steadier. A slight change in the position of the slider B on EF may be necessary during the first 50 or 100 hours, but finally it may be secured to the baseboard, and need only be altered if other valves are used, etc.

The valves should not be interchanged, used alone or together for any other purpose, or operated at values of the H.T. voltage other than the normal one employed during compensation, or the state of equilibrium will be disturbed. The L.T. battery should therefore be *connected last* and *disconnected first*, and the apparatus not operated if the H.T. circuit is broken by the omission of one or both of the anode resistances.

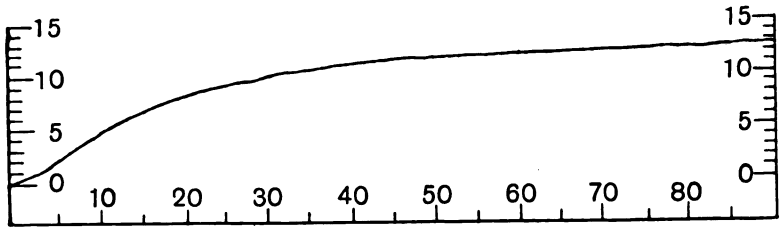
Trouble arising from electrostatic effects can be largely eliminated if it is observed that, as far as possible, the earth terminal of the apparatus is connected to the *same* earth wire as any other apparatus (i.e. ionisation chamber, battery, etc.) employed with it, and that if water, or other safety resistances are employed, they are placed near the ionisation chamber, and *not* in the earth lead.

Steadiness of the Zero.

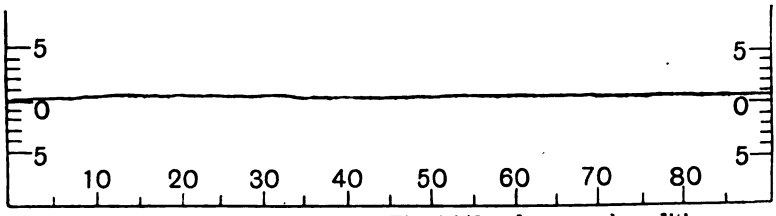
In order to investigate the steadiness of the zero of the system, and to compare it with those of other systems, a recording galvanometer, kindly lent by Dr C. D. Ellis, was used in place of the

Digitized by Google

ordinary one. This instrument, which gave a full scale deflection for 25 micro-amperes, enabled a great deal of information to be

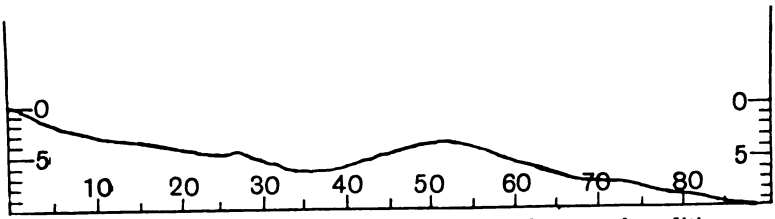


(*a*) Curve for compensated system [Fig. 1 (*c*)] before temperature equilibrium has been established.

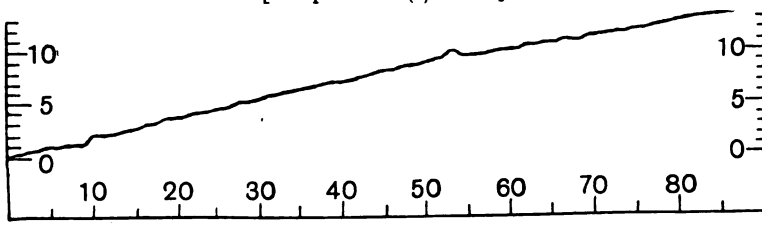


(*b*) Curve for compensated system [Fig. 1 (*c*)] under normal conditions.

Deflections in micro-amperes.



(*c*) Curve for uncompensated system [Fig. 1 (*a*)] under normal conditions. [Compare with (*b*) above.]



Time in minutes.

(*d*) Curve for uncompensated system [Fig. 1 (*b*)] under normal conditions. [Compare with (*b*) above.]

Fig. 4.

obtained which might otherwise have been overlooked, and greatly facilitated the investigation.

Digitized by Google

The recorder, arranged to take readings every half minute, was started immediately after the valves had been switched on and the system balanced, the grids being left unconnected. Except during the first few days, during which the system was establishing its state of equilibrium as previously explained, the curve traced by the recorder was of the form shown in Fig. 4(*a*), the same form being obtained on repetition. For the first half hour or hour, the zero drifts to a new position, the amount of drift in this case being about 12 micro-amperes. This drift is due to the fact that a temperature equilibrium has to be established between the various electrodes within the valve. If the system be re-balanced at the end of an hour after switching on, the zero is very steady, though a very small drift is almost always present, even after the system has been in operation for several hours. Generally, this does not exceed $\frac{1}{2}$ to 1 micro-ampere per hour, and takes place at a *steady rate*, so that allowance may be made for it by taking the initial and final zero positions during an experiment.

This residual drift is probably due, not to imperfect compensation, but to the fact that the physical conditions of the two filaments change slowly when in use, at slightly different rates. The drift is shown in the record of Fig. 4(*b*), taken several hours after first switching on. It will be observed how regular this effect is, there being little trace of any other disturbance present.

For comparison, two curves, Figs. 4(*c*) and 4(*d*), were obtained, when the recorder was connected to a single valve uncompensated system [Fig. 1(*a*)] and a double valve uncompensated system [Fig. 1(*b*)]. These curves were taken under the same conditions as the curve of Fig. 4(*b*), with which they should be compared. The superiority of the compensated system as a physical instrument is self-evident.

Characteristics of the System.

The behaviour of the instrument is very similar to that of a quadrant electrometer used in conjunction with a leak across the quadrants, the ionisation current in both cases being measured by the steady deflection it gives rise to, the leak in the case of the valve system being the natural grid leakage*. It will therefore be convenient and useful to effect a comparison between the two types of instruments, and to refer to the characteristics of the system (e.g. capacity, insulation, etc.) as if it were an electrometer used with a leak.

* The similarity is carried further by the fact that the perfectly insulated grid of a hard valve would retain a negative charge indefinitely. In such a case, a negative ionisation current to the grid would result in a steady increase in the potential of the latter, giving rise to a regular movement of the galvanometer. The current could, in this way, be measured in a similar manner to that employed for an electrometer without a leak. Such a system might, however, be unstable without a leak to earth of some sort.

Digitized by Google

Capacity. The electrostatic capacity of the grid on open circuit is usually of the order of a few centimetres, the system in this respect corresponding to an electrometer of low capacity.

Voltage sensitivity and uniformity of scale. By applying various potentials to the grid by means of a potentiometer connected to the negative end of the filament (and hence, to earth), and observing the galvanometer deflections, it was established that the deflection was proportional to the change in grid potential*. A change of grid potential of one milli-volt gave rise to a galvanometer current of 1/3·15 micro-ampere. Used in conjunction with a galvanometer of sensitivity 200 mm. per micro-ampere, the system has therefore a voltage sensitivity of 6350 mm. per volt. The normal "floating" potential of the grid was found to be slightly negative. This corresponds to contact potential in the electrometer.

Current sensitivity and uniformity of scale. For the purpose of determining the current sensitivity of the system, a direct comparison was made with a quadrant electrometer and leak. Various positive and negative ionisation currents were driven, in turn, to the electrometer, and on to one of the grids, and the deflections of the two instruments were compared. The deflections caused by currents up to $3{\cdot}5 \times 10^{-12}$ amperes were measured quantitatively, while the effect on the system of currents greater than these was determined qualitatively.

It was found that, in general, a given negative ionisation current produced a slightly larger deflection than a positive one of equal amount. The current amplification factor *increases with increasing negative grid potential, and decreases with increasing positive grid potential.* The increase in magnification becomes very large for large negative ionisation currents—a fact which can be utilised to exaggerate phenomena in lecture demonstrations. With the particular valves used, the current amplification factor for positive ionisation currents up to $3{\cdot}5 \times 10^{-12}$ amperes was found to be almost constant, and of the order of $1{\cdot}13 \times 10^{5}$†.

The reason for the variation of the current amplification factor for positive and negative currents lies in the difference in the rate of leakage from the grid of positive and negative charges. Grid leakage is due to (1) imperfect insulation of the grid, and (2) the emission current between the grid and the filament. While the grid current due to the former is a linear function of the grid potential—the insulation resistance of the grid being constant—

* This shows that the valve is operated on a suitable portion of its triode characteristic curve.

† As one of the two valves employed is sometimes found to be more sensitive than the other for current amplification, it is advisable to try both in turn and observe which is the better. The small differences in grid current characteristics and hardness, which affect the sensitivity, are not usually apparent in the tests to which valves are subjected by the makers.

Digitized by Google

the emission current increases with increasing positive grid potential, and decreases, tending to zero, with increasing negative grid potential. The grid leakage is therefore greater for a positive charge on the grid than for a negative, so that a given negative ionisation current raises the grid to a numerically higher potential than does a positive one of equal amount. The presence of a trace of gas in the valve, i.e. the valve not being perfectly hard, results in an increase of both positive and negative leakage currents, which tends to make the current amplification factor more nearly uniform for positive and negative ionisation currents.

Leak. A distinction must be made between the insulation resistance of the grid when the filament is cold, and the resistance between the grid and filament when the latter is emitting electrons.

The former, determined in the usual way for high resistances, i.e. condenser discharge, was found to be of the order of 70,000 to 85,000 megohms*.

The latter, which corresponds to the leak used with an electrometer to obtain steady deflections, was estimated by calculating, from the results of the two preceding sections, the rise of grid potential produced by unit ionisation current. This gave a value of 360 megohms, very much smaller than the cold insulation resistance. While therefore good insulation of the ionisation chamber is desirable, it is evidently not necessary to have such a high degree of insulation as when working with an electrometer, a very important advantage.

The Apparatus considered as an Electrometer.

Collecting the data given in the preceding section, it may be said that, in conjunction with a galvanometer of sensitivity 200 mm. per micro-ampere, the system can be regarded as equivalent to an electrometer of small capacity, having a voltage sensitivity of 6350 mm. per volt, and shunted by a leak of 360 megohms. The resistance of the leak increases with increasing negative current, and decreases with increasing positive current. In addition, it has the advantages of portability, compactness, and ease of erection and operation. The lower time of swing, dependent only upon the type of galvanometer employed, is also an advantage.

The lower limit to the minimum current that can be measured by the apparatus is governed by the steadiness of the zero. In spite of careful screening, it is almost impossible to eliminate completely small fluctuations of the galvanometer zero, due, not to imperfect compensation, but to electrical disturbances from induction coils, etc., which are picked up by the apparatus. For this particular apparatus, the lower limit appeared to lie between

* This includes leakage of the valve cap and valve holder, etc.

Digitized by Google

10^{-13} and 10^{-12} amperes, the deflections produced by currents less than 10^{-13} amperes being of the same order as the galvanometer zero fluctuations. Increasing the sensitivity of the galvanometer beyond a certain limit does not therefore extend the range. The current amplification factor could, of course, be increased, by giving to the grid a suitable negative bias through an ionisation leak*, but this would probably result in increased unsteadiness of the zero also.

The explanation of the unsteadiness caused by electrical disturbances is interesting. When an inductive circuit is broken, as when an induction coil is operated near the apparatus, an electric wave is set up, which induces an electric oscillation in the leads connected to the grid. The alternations of potential on the grid are rectified, in virtue of the dissymmetry of the grid leakage for positive and negative grid potentials, so that the grid acquires a mean negative potential, causing a reduction of the plate current. The characteristics of the two valves being different, a nett change takes place in the galvanometer current, resulting in a kick for each wave train picked up by the apparatus. Balance is restored as the grids recover their normal floating potential.

The similarity between this effect and the grid condenser and leak method of rectification suggests the possibility of using the apparatus as a sensitive wireless detector†.

The use of the instrument is not confined to the measurement of ionisation currents. By connecting a high resistance between the grid and earth, it can be used in the usual way for amplifying other small currents, *provided that the circuit contains sufficient potential to drive the current through the high resistance.* This will be the case when the other resistances in the circuit are large compared to the auxiliary resistance connected between the grid and earth. The current amplification factor for such a system is given by

$$\frac{\mu S}{2X + G\left(1 + \frac{X}{R}\right)},$$

S being the auxiliary resistance. When used for measuring ionisation currents, S takes the form of the natural grid leakage, viz. 360 megohms, corresponding to a value for the factor equal to $1{\cdot}15 \times 10^5$.

With a comparatively low value of S, which might have to be

* An ionisation or other saturation leak would have to be used to smooth out fluctuations of the bias battery voltage.

† This type of apparatus has been tried by Mr J. A. Ratcliffe for signal-strength measurements, and found to be very useful for this type of work, being superior to an ordinary valve detector on account of its steady zero, and less troublesome than a crystal rectifier.

Digitized by Google

used in certain circumstances, the value of the factor would be considerably reduced, a large current magnification only being obtained with an insulated grid.

In conclusion, it may be said that the method of compensation can be applied to other valve apparatus where a steady zero is required. By omitting the galvanometer G [Fig. 1(*c*)] the system becomes a voltage amplifier. A change E in the grid potential gives rise to a potential difference equal to $E\mu R/(X+R)$ between the two anodes, the value of the factor being as large as 30 or 40 when suitable "resistance-capacity" valves, having a large value of μ, are used in conjunction with anode resistances of the order of 10^6 ohms.

If either or both of the grids are to be earthed, through tuning coils, transformers, etc., *compensation should be effected under the conditions under which the apparatus will be operated*, and not with the grids insulated, as the characteristic curves (Fig. 2) are different in the two cases.

Summary.

A method of using a valve for amplifying ionisation currents 100,000 times is described, which avoids the instabilities usually associated with such apparatus. The necessary conditions to be satisfied, while extremely simple, are based on theoretical conditions which are discussed in full. Used in conjunction with a galvanometer of sensitivity 200 mm. per micro-ampere, the system behaves in a similar manner to a low capacity quadrant electrometer of sensitivity 6350 mm. per volt, shunted by a leak of 360 megohms, the value of the latter being slightly greater for negative currents than for positive. In addition, it has the advantages of portability, compactness, and ease and rapidity of erection and operation, being specially suitable for lecture demonstrations.

The method of compensation can also be applied to other valve circuits, resulting in a much steadier zero.

The experimental work in connection with the above was carried out at the Cavendish Laboratory, Cambridge, and the writer desires to thank Sir Ernest Rutherford for his encouragement, and for providing facilities for it.

He also desires to express his sincere thanks to Dr C. D. Ellis, for his interest in the work, and for valuable suggestions and advice tendered during the course of it. In particular, he is indebted to him for the temporary loan of a recording galvanometer, bought with a Government grant from the Royal Institution, which proved of great value.

Digitized by Google

The production of characteristic X-rays by electronic impact.
By L. H. Thomas, B.A., Trinity College.

[*Received* 23 May, *read* 25 July, 1927.]

Rosseland proposed a theory of the production of characteristic X-rays*. He calculated the number of atoms from which on the average an electron with definite energy will remove an electron from a particular core orbit, but neglected the velocity of the core electron in its orbit and the increased velocity of the impinging electron due to the atomic field. Taking these into account considerably alters his formula.

Suppose that an electron, charge e and mass m, moves with velocity V through a cloud of N electrons per unit volume moving in random directions with velocity v. Then it can be shown that the chance of a collision in time dt in which the first electron loses energy between q' and $q'+dq'$ and momentum between s and $s+ds$ is

$$\frac{4\pi Ne^4}{vVs^4}\,dt\,ds\,dq' \qquad \ldots\ldots(1),$$

if such a collision is possible at all†. If V', v' are the velocities after the collision, so that

$$q' = \tfrac{1}{2}m(v'^2 - v^2) = \tfrac{1}{2}m(V^2 - V'^2) \qquad \ldots\ldots(2),$$

the conditions that such a collision is possible are

$$m\,|\,v'-v\,| < s < m(v'+v),$$
$$m\,|\,V-V'\,| < s < m(V+V').$$

Now suppose $q'>0$ and $V'>v$, so that, from (2), $V>v'$ and the limits for s are $m(v'-v)$ and $m(v'+v)$. Hence, writing $\frac{1}{2}mv^2 = k_2$, and integrating (1) between these limits, the chance that the first electron loses energy between q' and $q'+dq'$ is

$$\frac{2\pi Ne^4}{mV}\left[\frac{1}{q'^2} + \frac{4}{3}\frac{k_2}{q'^3}\right]dq'\,dt \qquad \ldots\ldots(3)‡.$$

Now consider the removal from an atom of an electron with kinetic energy in its orbit k_2 and ionisation energy w, by an electron with kinetic energy q_1. In order to approach the former, the latter must gain kinetic energy from the atomic field, and just before collision will have kinetic energy $k_1 = q_1 + w + k_2$. Formula (3) can be applied, noting that dt is to be replaced by dx/V, where

* Rosseland, *Phil. Mag.*, Vol. xlv

† Cf. *Proc. Camb. Phil. Soc.*, Vol.

‡ Cf. Williams, *Manchester Memo*
formula for this chance is obtained.

Digitized by Google

$\frac{1}{2}mV^2 = k_1$ and dx represents distance travelled by the first electron (dt should really be replaced by dx/v_1, where $\frac{1}{2}mv_1^2 = q_1$, but the chance so obtained must be decreased in the ratio v_1/V as the space-density of electrons with given energy is decreased in this way when a field increases their velocity). Hence the chance that in path distance dx an electron with energy q_1 will lose energy q' to an atomic electron with ionisation energy w and kinetic energy k_2—there being N such electrons per unit volume—is

$$\frac{N\pi e^4 dx}{(q_1 + w + k_2)}\left[\frac{1}{q'^2} + \frac{4}{3}\frac{k_2}{q'^3}\right]dq' \qquad \text{.....(4).}$$

For the application to the excitation of X-rays, suppose that the material (metal of the anti-cathode) has known stopping-power S for electrons, defined by

$$-\frac{d(v_1^4)}{dx} = \frac{8\pi Ne^4}{m^2}S \qquad \text{.....(5).}$$

The largest part of S is due to collisions with electrons of small w, and varies nearly as $\log q_1$ so long as q_1 is large, while it is only for large q_1 that an electron can be removed from a core orbit at all. Thus, if the excess voltage of the impinging electrons over the critical voltage is not large compared with the latter, S may be considered constant in finding the total number removed. For path length dx the number will be

$$\frac{N\pi e^4 dx}{(q_1 + w + k_2)}\int_w^{q_1}\left[\frac{1}{q'^2} + \frac{4}{3}\frac{k_2}{q'^3}\right]dq'$$
$$= \frac{N\pi e^4 dx}{(q_1 + w + k_2)}\left[\frac{1}{w} - \frac{1}{q_1} + \frac{2}{3}\frac{k_2}{w^2} - \frac{2}{3}\frac{k_2}{q_1^2}\right].$$

From (5), $$dx = -\frac{q_1\,dq_1}{N\pi e^4 S},$$

so that the total number of electrons removed from such a level by one impinging electron of original energy q is

$$\frac{1}{S}\int_w^q \frac{q_1}{(q_1 + w + k_2)}\left[\frac{1}{w} - \frac{1}{q_1} + \frac{2}{3}\frac{k_2}{w^2} - \frac{2}{3}\frac{k_2}{q_1^2}\right]dq_1,$$

i.e. $$\frac{1}{3S}\left[(q-w)\left(\frac{2k_2}{w^2} + \frac{3}{w}\right) - \frac{2k_2}{(k_2 + w)}\log\left\{\frac{q(2w + k_2)}{w(q + w + k_2)}\right\}\right.$$
$$\left. - \left\{3 + (k_2 + w)\left(\frac{2k_2}{w^2} + \frac{3}{w}\right)\right\}\log\left\{\frac{q + w + k_2}{2w + k_2}\right\}\right] \qquad \text{.....(6).}$$

The expression (6) multiplied by S should replace the expression

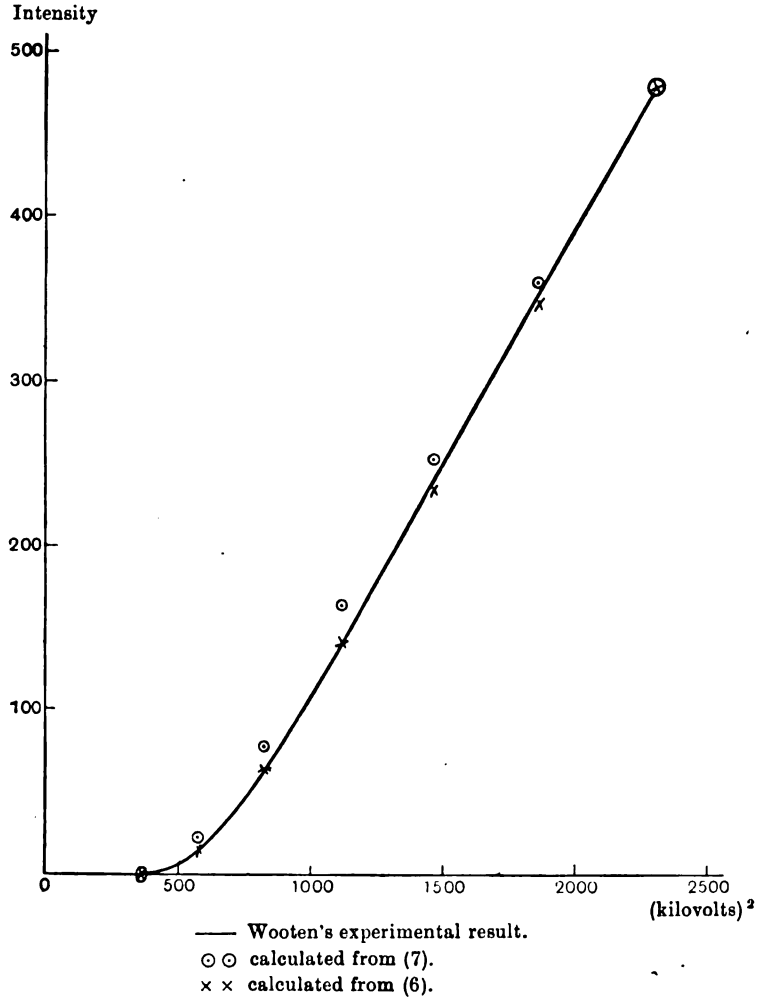
$$\left\{\frac{E_0 - W}{W} - \log\frac{E_0}{W}\right\} \qquad \text{.....(7)}$$

in Rosseland's paper*, where E_0, W correspond to q, w.

* P. 78, *loc. cit.*

Digitized by Google

In the following graph the variations in the intensity of X-rays with applied voltage calculated from (6) and (7) are compared with those found by Wooten* for the $K\alpha$ line of molybdenum. For the K-level of molybdenum, $w = 19{\cdot}2$ kilovolts, $k_2/w = 1{\cdot}197$. The absolute values were adjusted to fit at intensity 480 for $q = 48$ kilovolts.



In conclusion I wish to thank Dr Nishina and Dr Ray for suggesting this calculation, and Professor Bohr for his encouragement and help.

* Wooten, *Phys. Rev.*, Vol. XIII (1919), p. 76.

Digitized by Google

On the Effect of Continued Small Additions of Poisonous Substances on the Velocity of Gaseous Catalytic Reaction in Closed Vessels. By Dr F. HURN CONSTABLE, Fellow of St John's College.

[*Received* and *read* 2 May 1927.]

The Langmuir-Frenkel theory of surface action has been shown to involve a simple relation between the partial pressure of the catalyst poison, and the reduction in reaction velocity which results*. Very interesting observations have been made by Beebe† showing that the heat of adsorption of carbon monoxide on a reduced copper catalyst falls considerably and reaches a limiting value as the surface approaches saturation. Pease and Stewart‡ have shown that the presence of less than one per cent. of the total carbon monoxide that a given copper catalyst can adsorb will reduce the reaction velocity of the combination of hydrogen and ethylene at the surface of the copper by 88 per cent.

It is interesting to note that there are a number of explanations of the phenomenon noticed by Beebe, and that it would be expected to occur to some extent even with a homogeneous surface.

(*a*) As the surface becomes increasingly covered by the adsorbed molecules the work of desorption falls, because the external fields from the surface are weakened§.

(*b*) The portions of the surface with the most intense fields adsorb gas at pressures low enough for the remainder of the surface to be inactive as an adsorbent.

(*c*) Multimolecular films may be formed, the work of desorption being very much greater for the first unimolecular film, and falling to a limiting value as the number increases.

(*d*) Clusters of gas molecules may adhere to particular spots on the surface; thus though the surface may adsorb far less gas than would be required to form a complete unimolecular layer, it behaves as if unimolecular layers had been formed.

(*e*) The surface is covered with unimolecular layers interspersed with portions of the surface that are bare; thus the whole surface is incapable of holding gas in the adsorbed state.

* Constable, *Proc. Camb. Phil. Soc.*, vol. 23, pp. 172–182 (1926), "On the behaviour of the centres of activity of saturated surfaces during the initial stages of unimolecular reactions"; pp. 593–606 (1926), "Surface adsorption and the velocity of chemical action at gas solid interfaces."

† *J. Phys. Chem.*, vol. 30, pp. 1538–1544 (1926).

‡ *J. Amer. Chem. Soc.*, vol. 47, p. 1235 (1925).

§ Becker, *Phys. Rev.*, vol. 28, pp. 341–361, has shown by direct experiment with caesium films on tungsten that the work necessary to remove an atom of caesium decreases as the fraction of the surface covered increases; but that the work done in removing a caesium ion increases under similar conditions.

Digitized by Google

The observations of Pease and Stewart show that the same small quantity of poison added to a closed vessel in which catalysis is occurring has a vastly greater effect initially than when the pressure of the poison has become considerable. The theory of centres of activity developed by Pease, Taylor, and Constable implies that there is a distribution of centres of all degrees of activity over the catalyst surface, and it is interesting to examine the data of Pease and Stewart to see whether any information can be obtained from their results as to the type of distribution present.

The direct result of the Langmuir-Frenkel theory is that saturated surfaces over which there is no variation of the activity of the individual centres should be much more sensitive to the first small quantities of poison than to the amounts which follow; and the purpose of this paper is to show the nature of this variation in the case of uni- and bi-molecular reactions. It can then be seen that the data on the poisoning of the hydrogenation of ethylene by increasing quantities of carbon monoxide show definitely that the mean life of the carbon monoxide molecule is very much greater on the centres which are most active in hydrogenation than on the rest of the surface, thus bringing forward additional evidence in favour of the enhanced fields on the most active spots on the surface.

The Initial Stages of the Poisoning of Unimolecular and Bimolecular Gas Reactions in Closed Vessels.

In these cases it is experimentally convenient to keep the pressure of the reactants constant, and measure the decrease in the initial catalytic activity as the absolute pressure of the poison increases. The initial rate of bimolecular reaction can conveniently be measured by the initial rate of fall of pressure in the closed vessel. Let the reaction be

$$A + B \rightarrow C,$$

and let a = the initial pressure of the reactant A,
b = the initial pressure of the reactant B,
x = the pressure of the poison.

Then, following the method of calculation given by the author*, the fraction of any time interval during which the centres of activity that activate A are occupied by A is

$$\frac{\mu_A \tau_A a}{\mu_A \tau_A a + \mu_B \tau_B b + \mu_C \tau_C x + \xi},$$

* *Nature*, vol. 117, p. 230 (1926). The symbols are defined here.

Digitized by Google

where ξ is the fraction of any time interval that the centres holding A are left free,

$$\mu = \left(\frac{1}{2\pi MRT}\right)^{\frac{1}{2}},$$

and τ_L is the mean life of the molecule L on the centre. Similarly, the fraction of any period of time that the second series of centres are occupied by B is

$$\frac{\mu_B \tau_B' b}{\mu_A \tau_A' a + \mu_B \tau_B' b + \mu_C \tau_C' x + \xi'},$$

for the surface fields of force holding and activating the molecules B need not be the same as those holding A.

Bimolecular reaction is impossible while the reactants are separated*; thus the velocity is proportional to the probability that the reactants should be present side by side simultaneously on the centres of activity capable of activating both. The rate of reaction, conveniently measured by the initial rate of fall of pressure, is given by

$$\frac{da}{dt} = \frac{\mu_A \tau_A \mu_B \tau_B' ab}{(\mu_A \tau_A a + \mu_B \tau_B b + \mu_C \tau_C x + \xi)(\mu_A \tau_A' a + \mu_B \tau_B' b + \mu_C \tau_C' x + \xi')}.$$

The fractional reaction velocity is given by the ratio of the initial velocity, when the pressure of poison is x, to the initial rate of reaction in the absence of poison.

Therefore

$$v = \frac{1}{\left(1 + \left\{\frac{\mu_C \tau_C}{\mu_A \tau_A a + \mu_B \tau_B b + \xi}\right\} x\right)} \frac{1}{\left(1 + \left\{\frac{\mu_C \tau_C'}{\mu_A \tau_A' a + \mu_B \tau_B' b + \xi'}\right\} x\right)} \quad \ldots(1)$$

or

$$v = \frac{1}{(1 + \lambda_1 x)} \frac{1}{(1 + \lambda_2 x)}.$$

In some bimolecular reactions the poison may only have an appreciable mean life on one type of centre; thus the expression obtained for the reduction of the rate of the initial stages of the combination is the same as would be obtained for a unimolecular reaction, *i.e.*,

$$v = \frac{1}{\left(1 + \left\{\frac{\mu_C \tau_C}{\mu_A \tau_A a + \mu_B \tau_B b + \xi}\right\} x\right)} \quad \ldots\ldots\ldots\ldots(2).$$

* Larmor, *Proc. Lit. Phil. Soc. Manchester*, vol. 52 (No. 10), p. 28 (1908), has pointed out the possibilities arising from this principle in connection with gas reactions.

Digitized by Google

The Efficiency of Continued Small Additions of Poison in Reducing the Reaction Velocity in the Initial Stages of Catalysis in Closed Vessels.

The percentage reduction in the initial rate of reaction due to the addition of poison at pressure x is given by $100\,\delta v/v$; and a measure of the effectiveness of the next increment δx in the pressure of the poison is the value of $\delta v/v\,\delta x$.

For bimolecular reaction

$$\frac{1}{v} = 1 + \alpha x + \beta x^2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3),$$

where $$\alpha = \frac{\mu_C \tau_C}{\Sigma \mu_A \tau_A a + \xi} + \frac{\mu_C \tau_C'}{\Sigma \mu_A \tau_A' a + \xi'} \equiv \lambda_1 + \lambda_2,$$

and $$\beta = \frac{\mu_C \tau_C \mu_C \tau_C'}{(\Sigma \mu_A \tau_A a + \xi)(\Sigma \mu_A \tau_A' a + \xi')} \equiv \lambda_1 \lambda_2;$$

therefore $$\frac{1}{v}\frac{\partial v}{\partial x} = -\alpha \left\{ \frac{1 + \frac{2\beta}{\alpha} x}{1 + x(\alpha + \beta x)} \right\} \ldots\ldots\ldots\ldots\ldots\ldots (4).$$

For unimolecular reaction

$$\frac{1}{v} = 1 + \alpha x,$$

where $$\alpha = \frac{\mu_O \tau_O}{(\mu_A \tau_A a + \mu_B \tau_B b + \xi)};$$

thus $$\frac{1}{v}\frac{\partial v}{\partial x} = -\alpha \left\{ \frac{1}{1 + \alpha x} \right\} \ldots\ldots\ldots\ldots\ldots\ldots (5).$$

Since α is the sum of λ_1 and λ_2, while β is the product $\lambda_1 \lambda_2$, the maximum value of $\frac{1}{v}\frac{\partial v}{\partial x}$ is $-\alpha$ and this value falls considerably as x increases* since the multiplying factor becomes continuously

* Note that $$\frac{1 + \frac{2\beta}{\alpha} x}{1 + x(\alpha + \beta x)} \equiv \frac{1 + \frac{2\lambda_1 \lambda_2}{\lambda_1 + \lambda_2} x}{1 + x\left(\lambda_1 + \lambda_2 + \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2} x\right)},$$

and is always less than unity provided that

$$(\lambda_1 + \lambda_2) + \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2} x > \frac{2\lambda_1 \lambda_2}{\lambda_1 + \lambda_2},$$

i.e.

$$(\lambda_1 + \lambda_2)^2 + \lambda_1 \lambda_2 x > 2\lambda_1 \lambda_2,$$

$$\lambda_1^2 + \lambda_2^2 + \lambda_1 \lambda_2 x > 0,$$

which is always satisfied by a molecule which has a positive mean life on the centres of activity. The variation in ξ with x has been neglected, but such variation would only intensify the effect already predicted. The variation is absent with saturated centres, and it seems probable that the 'centres of activity' in many cases are always covered by adsorbed molecules.

Digitized by Google

smaller; thus equation (4) shows that for bimolecular reaction the efficiency of a catalyst poison, on a catalyst composed of a uniform distribution of centres of activity, falls off as the pressure of the poison increases. Such evidence alone is thus valueless to show the presence of a distribution of centres of activity. The relation for unimolecular reaction shows the fall in the efficiency still more plainly, and it is probable that quite a number of bimolecular reactions follow this simple law.

Thus, in the simple unimolecular formula (2), unless it be shown definitely that λ is variable with x, it cannot be deduced from evidence of this type that there is a distribution of centres of catalytic activity over the surface.

Pease and Stewart have made deductions of this nature from observations of the pressure of carbon monoxide and the corresponding reduction in the rate of hydrogenation of ethylene by copper in a closed vessel. Their results show that λ is variable in formula (2).

TABLE I. *Showing the variation of* λ *with* x.

Volume of carbon monoxide in catalyst bulb, i.e. kx where k is constant	Fractional activity of catalyst	$\frac{1/v-1}{x}=1/k\lambda$
0·05	0·108	165
0·08	0·092	123
0·33	0·084	33
0·69	0·077	17
1·97	0·050	9·6
9·14	0·013	8·5

The interpretation to be placed on the variation in λ is that the mean life of the carbon monoxide molecule relative to the reactants is many times greater on the most active centres than it is on the least active. If $1/v$ be plotted against x, these data give a straight line which cuts the x axis at a value of $x = 0{\cdot}88$. The behaviour of the surface therefore approximates to that of a homogeneous distribution of centres once the most active portions of the surface are covered. The results below are calculated assuming the initial activity to be 0·12. They show an agreement with the values to be expected for homogeneous distribution which is good considering the experimental errors in the measurement of the pressure of the poison.

Digitized by Google

TABLE II. *Showing the approximate constancy of* λ *when the most active centres are poisoned.*

Fractional reaction velocity	Calculated fractional reaction velocity from $1/v = 1 + 0{\cdot}88x$	x
1·00	1·00	0·00
0·90	0·96	0·05
0·77	0·93	0·08
0·70	0·79	0·33
0·64	0·62	0·69
0·42	0·36	1·97
0·11	0·11	9·14

The ratio of the greatest value of $k\lambda$ in Table I (165) to the value in Table II (0·88) shows that the carbon monoxide molecule occupies the most active centres very much longer than the reacting gases; while on the less active centres the mean life of the carbon monoxide molecules is comparable with that of the reactive gases. The observations of Pease and Stewart thus show clearly that a great difference exists between the surface fields of force on the centres of activity involved in the two cases.

Digitized by Google

A Mathematical Theory of Natural and Artificial Selection, Part V: Selection and Mutation. By Mr J. B. S. HALDANE, Trinity College.

[*Received* 21 May, *read* 25 July, 1927.]

New factors arise in a species by the process of mutation. The frequency of mutation is generally small, but it seems probable that it can sometimes be increased by changes in the environment (1, 2). On the whole mutants recessive to the normal type occur more commonly than dominants. The frequency of a given type of mutation varies, but for some factors in Drosophila it must be less than 10^{-6}, and is much less in some human cases. We shall first consider initial conditions, when only a few of the new type exist as the result of a single mutation; and then the course of events in a population where the new factor is present in such numbers as to be in no danger of extinction by mere bad luck. In the first section the treatment of Fisher (3) is followed.

In a large population let p_r be the chance that a factor present in a zygote at a given stage in the life-cycle will appear in r of its children in the next generation. If the individual considered is homozygous, this is the chance of leaving r children, if mutation is neglected. Let $\sum_{r=0}^{\infty} p_r x^r = f(x)$. Therefore $f(1) = 1$, $f(0) = p_0$, the probability of the factor disappearing, while $f'(1) = \sum_{r=0}^{\infty} r p_r$, *i.e.*, the probable number of individuals possessing the factor in the next generation. The probability of m individuals bearing one each of the factors considered leaving r descendants is clearly the coefficient of x^r in $[f(x)]^m$, if we neglect the possibility of a mating between two such individuals, which we may legitimately do if m is small compared with the total number of the population. If then the probability of the factor being present in r zygotes of the nth generation be the coefficient of x^r in $F(x)$, the corresponding probability in the $(n+1)$th generation is the same coefficient in $F[f(x)]$. Hence if a single factor appears in one zygote, the probability of its presence in r zygotes after n generations is the coefficient of x^r in $S_f^n(x)$, *i.e.* $f(f(f \ldots f(x) \ldots))$, the operation being repeated n times. The probability of its disappearance is therefore $\lim_{n \to \infty} S_f^n(0)$. By Koenigs' theorem (4) this is the root of $x = f(x)$ in the neighbourhood of zero.

Digitized by Google

Now in the case of a dominant factor appearing in a population in equilibrium, and conferring an advantage measured by k, as in Part I (5), $f'(1) = 1 + k$. Since $f'(x)$ and $f''(x)$ are positive when x is positive, and $f(0)$ is positive, $x = f(x)$ has two and only two real positive roots, one equal to unity, the other lying between 0 and 1, but near the latter value if k be small. Hence any advantageous dominant factor which has once appeared has a finite chance of survival, however large the total population may be.

If a large number of offspring is possible, as in most organisms, the series p_n approximates to a Poisson series, provided that adult organisms are counted, and since $f'(1) = 1 + k$, $f(x) = e^{(1+k)(x-1)}$. Hence the probability of extinction $1 - y$ is given by

$$1 - y = e^{-(1+k)y}.$$

Hence
$$(1 + k)y = -\log(1 - y) \quad \ldots\ldots\ldots\ldots\ldots\ldots (1{\cdot}0),$$

and
$$k = \frac{y}{2} + \frac{y^2}{3} + \frac{y^3}{4} + \ldots,$$

and if k be small, $y = 2k$ approximately. Hence an advantageous dominant gene has a probability $2k$ of survival after only a single appearance in an adult zygote, and if in the whole history of a species it appears more than $\frac{\log_e 2}{2k}$ times it will probably spread through the species. But, however large k may be, the factor may be extinguished after a single appearance. Thus if $k = 1$, so that the new type probably leaves twice as many offspring as the normal, the probability of its extinction is still ·203. If in any generation there are m dominant individuals the probability of extinction is reduced to y^m, where y is the smaller positive root of $x = f(x)$. When k is small this reduces to $(1 - 2k)^m$. Hence if in any generation more than $\frac{\log_e 2}{2k}$ adult dominants exist, the factor will probably spread through the whole population.

On the other hand a recessive factor whose phenotype is advantageous has a quite negligible advantage in a random mating population provided that the number of its bearers is small compared with the square root of the total population. This is best seen by considering the case of a hermaphrodite: in a dioecious organism the argument, though similar, is more complicated. Let N be the fixed number of the population, and z_n the number of heterozygotes plus double the number of recessives for the factor A in the nth generation. It therefore produces gametes in the ratio $(2N - z_n)A : z_n a$. If now the recessives have a small advantage measured by k, the probabilities of production of each genotype in the next generation are

$$(2N - z_n)^2 AA : 2z_n(2N - z_n)Aa : (1 + k)z_n^2 aa.$$

Digitized by Google

Hence if, as above, $f(x)$ be the function defining the probable number of offspring of a dominant, so that $f'(1) = 1$, the probability of r heterozygotes in the $(n+1)$th generation is the coefficient of r in

$$[f(x)]^{\frac{2Nz_n(2N-z_n)}{4N^2+kz_n^2}},$$

that of r recessives the same coefficient in

$$[f(x)]^{\frac{N(1+k)z_n^2}{4N^2+kz_n^2}}.$$

Hence the probability of z_{n+1} in the next generation is the coefficient of $x^{z_{n+1}}$ in

$$[f(x)]^{\frac{2Nz_n(2N+kz_n)}{4N^2+kz_n^2}},$$

or, approximately, if z_n be small compared with N, in

$$[f(x)]^{z_n\left(1+\frac{kz_n}{2N}\right)}.$$

The corresponding expression for a dominant factor is

$$[f(x)]^{z_n(1+k)}.$$

Hence provided that z_n is small the probability of escaping extinction is much smaller than k. I have been unable to evaluate it exactly, but it seems from a comparison with the case of a dominant factor, that the value of z_n such that the factor is as likely to survive as to be extinguished, is of the order of $\left(\frac{N}{k}\right)^{\frac{1}{2}}$, *i.e.* generally $>N^{\frac{1}{2}}$. So if N is sufficiently large the probability of a single mutation leading to the establishment of a recessive factor is negligible.

When the population is wholly self-fertilized or inbred by brother-sister mating, on the other hand, a recessive factor has almost as good a chance of survival as a dominant. With partial self-fertilization or inbreeding it can be shown by methods similar to those of Part II (6) that an advantageous recessive factor has a finite chance of establishment after one appearance, however large be the population.

If mutation occurs with a finite frequency any advantageous or not too disadvantageous factor will certainly be established. Consider a random mating population in which, in each generation, a proportion p of the A genes mutate to a, a proportion q of the a genes to A, and the coefficient of selection is k. Let u_n be the gametic ratio of the nth generation. But for mutation we should have

$$u_{n+1} = \frac{u_n(u_n+1)}{u_n+1-k};$$

Digitized by Google

allowing for mutation

$$u_{n+1} = \frac{(1-p)(u_n^2+u_n)+q(u_n+1-k)}{(1-q)(u_n+1-k)+p(u_n^2+u_n)}.$$

Hence $$\Delta u_n = \frac{ku_n^2 - pu_n(u_n+1)^2 + q(u_n+1)(u_n+1-k)}{u_n+1-k+pu_n(u_n+1)-q(u_n+1-k)} \quad \ldots(2{\cdot}0)$$

$$= \frac{ku_n}{u_n+1} - pu_n(u_n+1)+q(u_n+1) \quad \ldots\ldots\ldots\ldots\ldots(2{\cdot}1)$$

approximately, if p, q, and k are small, as is generally the case. It is clear that u_n must lie between $\frac{1-p}{p}$ and $\frac{q}{1-q}$, *i.e.* between $\frac{1}{p}$ and q approximately, and that when near these values it alters rapidly. But as p and q may be less than 10^{-6} these limits are very wide. The population is in equilibrium when

$$pu^3+(2p-q)u^2+(p-2q-k+kq)u-q+kq=0.$$

There is always one real positive root since p and q are positive and less than unity. If k be positive there is only one such root, defining a stable condition towards which the population tends when dominants have the advantage. If k or q be large compared with p this root approximates to $\left(\frac{k}{p}\right)^{\frac{1}{2}}$ or $\frac{q}{p}$ as the case may be, *i.e.* recessives nearly disappear. If p be of the same order of magnitude as the larger of k and q, u has a moderate value and the population is dimorphic. If p be much larger than k or q, u is small and approximates to $\frac{q}{p}$, *i.e.* dominants are rare.

If k be negative all the roots are positive if they are real, provided $q > 2p$ and $-k(1-q) > 2q-p$. They are real if $\frac{\Delta}{3k}$, *i.e.*

$$4(p+q)^3+[-27p^2+18p(p+q)(1-q)+(p+q)^2(1-q)^2]k+4p(1-q)^3k^2$$

is positive, that is to say, when q is small, if

$$4pk^2+(-8p^2+20pq+q^2)k+4(p+q)^2$$

is positive. All these three conditions can rarely be fulfilled, but such cases may presumably occur. Thus if $p = {\cdot}000{,}001$, $q = {\cdot}0004$, $k = -{\cdot}008$; $u^3-398u^2+7197{\cdot}8u-403{\cdot}2=0$. Therefore $u = {\cdot}057$, $18{\cdot}93$, or $379{\cdot}0$, giving $89{\cdot}5\,\%$, $0{\cdot}252\,\%$, or ${\cdot}000{,}693\,\%$ of recessives. In such a case the middle root defines an unstable equilibrium, the other two equilibria being stable. Thus the above considered population would be stable with only about seven recessives per million, the small tendency of dominant genes to mutate to recessive being balanced by reverse mutation. But if

Digitized by Google

a group containing more than one recessive gene in twenty were isolated from it, selection would be effective, and it would pass into a condition where only 10·5 °/₀ were dominants, this number being kept up by mutation.

Usually when k is negative there is only one real root. If p or $-k$ be large compared with q, it is small and approximates to $\frac{q}{p}$ or $\frac{-q}{k}$ as the case may be, so that dominants are rare. If q be of the same order of magnitude as the larger of p and $-k$, the root has a moderate value and the population is dimorphic. If q be larger than p or $-k$, u is large and approximates to $\frac{q}{p}$, so that recessives are few.

The rate of approach to equilibrium is given by

$$\frac{du_n}{dn} = \frac{ku_n}{u_n+1} - pu_n(u_n+1) + q(u_n+1) \quad \ldots\ldots(2{\cdot}2),$$

provided that the constants are small. The exact expression for n in terms of u_n depends on the nature of the roots and the side from which an equilibrium is being approached, but it always contains logarithmic terms. Hence the numbers of the rarer type of the population in succeeding generations always lie between two geometric series until equilibrium is nearly reached. That is to say, the march of events is comparatively rapid.

In a self-fertilizing population we can similarly show that

$$\Delta u_n = ku_n - pu_n(u_n+1) + q(u_n+1) \quad \ldots\ldots\ldots(2{\cdot}3).$$

Only one equilibrium is possible, and the course of events can readily be calculated in any given case. Similarly for a sex-linked factor

$$\Delta u_n = \frac{ku_n(u_n+3)}{3(u_n+1)} - pu_n(u_n+1) + q(u_n+1) \ldots\ldots(2{\cdot}4).$$

In this case if k be negative, three equilibria are sometimes found, and selection is more effective than in the autosomal case when recessives are rare.

To sum up, if selection acts against mutation, it is ineffective provided that the rate of mutation is greater than the coefficient of selection. Moreover, mutation is quite effective where selection is not, namely in causing an increase of recessives where these are rare. It is also more effective than selection in weeding out rare recessives provided that it is not balanced by back mutation of dominants. Mutation therefore determines the course of evolution as regards factors of negligible advantage or disadvantage to the species. It can only lead to results of importance when its frequency becomes large.

Digitized by Google

Addendum. Equilibrium and selection in Sciara and similar animals.

In Part I of this series all the then known types of single-factor Mendelian inheritance were discussed. Since then Metz (7) has discovered a new type in Sciara which is here treated on the lines of Part I. Gametogenesis is normal in the female, but spermatozoa are formed from maternal chromatin only. Hence there are two types of heterozygous male, which may be symbolized by $A(a)$ and $(A)a$ according as the A is received from the mother or father. They yield A and a spermatozoa respectively, the other genotypes behaving normally.

In the absence of selection let eggs and spermatozoa be produced by the mth generation in the proportions $u_m A : 1a$ and $v_m A : 1a$, respectively. The next generation is therefore:

$$\text{♀ } u_m v_m AA : (u_m + v_m)Aa : 1aa.$$

$$\text{♂ } u_m v_m AA : u_m A(a) : v_m (A)a : 1aa.$$

Hence

$$\left.\begin{aligned} u_{m+1} &= \frac{2u_m v_m + u_m + v_m}{u_m + v_m + 2} \\ v_{m+1} &= u_m \end{aligned}\right\} \quad \text{.....(3·0),}$$

which is the same as equation (6·0) of Part I (5). Hence, as in the above equation, we find, if y_m be the proportion of recessives in the mth generation,

$$y_m = y_\infty - \left(\frac{-1}{2}\right)^m c^{\frac{1}{2}} y_\infty^{\frac{1}{2}} + \left(\frac{-1}{2}\right)^{2m-1} c \quad \text{.....(3·1),}$$

where c is a constant depending on the initial conditions. Hence equilibrium is rapidly approached, the values in successive generations being alternately greater and less than the final value.

If selection occurs with a coefficient k in ♀s, l in ♂s, then

$$\left.\begin{aligned} u_{n+1} &= \frac{2u_n v_n + u_n + v_n}{u_n + v_n + 2 - 2k} \\ v_{n+1} &= \frac{u_n(v_n + 1)}{v_n + 1 - l} \end{aligned}\right\} \quad \text{.....(3·2).}$$

If the population is nearly in equilibrium apart from selection and k and l are small, so that u_n and v_n are nearly equal,

$$\Delta u_n = \frac{v_n - u_n}{2} + \frac{k u_n}{u_n + 1},$$

$$\Delta v_n = u_n - v_n + \frac{l u_n}{u_n + 1}, \text{ both approximately.}$$

Digitized by Google

Hence $$\Delta u_n = \frac{2k+l}{3} \frac{u_n}{u_n+1},$$

and $$\frac{2k+l}{3} n = u_n - u_0 + \log_e \left(\frac{u_n}{u_0}\right) \quad \text{.....(3·3),}$$

approximately. Selection therefore occurs much as with a normally inherited autosomal factor.

REFERENCES.

1. Harrison and Garrett. *Proc. Roy. Soc.* B, 99, p. 241, 1926.
2. Gager and Blakeslee. *Proc. Nat. Ac. Sci.* 13, p. 75, 1927.
3. Fisher. *Proc. Roy. Soc. Edin.* 42, p. 321, 1922.
4. Koenigs. *Darb. Bull.* (2) 7, p. 340, 1883.
5. Haldane. *Trans. Camb. Phil. Soc.* 23, p. 19, 1924.
6. Haldane. *Biol. Proc. Camb. Phil. Soc.* 1, p. 158, 1924.
7. Metz. *Proc. Nat. Ac. Sci.* 12, p. 690, 1926.

Digitized by Google

Digitized by Google

Digitized by Google

PROCEEDINGS
OF THE
Cambridge Philosophical Society.

A theorem of Dr P. Zeeman. By Mr J. H. GRACE, Peterhouse.

[*Received* 23 May, *read* 25 July 1927.]

1. The following theorem, first enunciated by Dr Zeeman in 1899*, has attracted the attention of several writers:

If five points A, B, C, D, E in three dimensions are such that the feet of the perpendiculars from E on the faces of the tetrahedron $ABCD$ are coplanar the relation between the five points is symmetrical.

It appears to me nevertheless that the true basis of the theorem has perhaps not been reached and that possibly the ensuing remarks may have some interest.

To begin with, it may be well to observe that the theorem belongs essentially to Euclidean space, though it is true in any number of dimensions, as is established in an elaborate investigation by Dr W. F. Meyer. I have therefore for the most part not troubled with general projective space. In the course of a proof of the theorem Dr W. Mantel showed by elementary reasoning that if we introduce five other points A', B', C', D', E' such that E, E', for example, are foci of a quadric of revolution inscribed in the tetrahedron $ABCD$, then D', E' are images in the plane ABC*. This theorem, from which Zeeman's can be readily deduced, is true for projective space, as will be seen (§ 6). Finally I must refer to a paper by Mr Richmond† wherein it is proved that, if A, B, C, D, E are any five points on a twisted cubic which cuts the plane at infinity in three points forming a triangle circumscribed to the imaginary circle at infinity, these points have what may be called the Zeeman property; he also establishes a converse theorem and extends the whole to n dimensions.

I shall work throughout in three dimensions, though putting the argument in such a form that it is readily extended to any number. For convenience Mr Richmond's theorem mentioned above is first established independently of Zeeman's result.

It may be as well to state that my main thesis is to establish the following theorem:

The section of the space figure $ABCDE$ by the plane at infinity is a system of ten points and ten lines arising from two triangles

* *Wiskundige Opgaben*, 1899–1902. Other references may be found in Mr Richmond's paper from which I take this.

† *Proc. C.P.S.*, vol. XXII, 1923, p. 34. The analogue of Mantel's theorem in two dimensions was known earlier and occurs in graphical statics. Obscurities in the present exposition will be partially removed if the reader will construct the corresponding argument in two dimensions.

Digitized by Google

in perspective; a conic Σ exists such that each of the points is the pole of one of the lines and the five original points have the Zeeman property if Σ is so related to the imaginary circle at infinity Γ that triangles can be inscribed in Σ which are self-conjugate for Γ.

2. Suppose that a twisted cubic through five points A, B, C, D, E cuts the plane at infinity in three points P, Q, R forming a triangle whose sides touch Γ, then I say that the feet of the perpendiculars from E on the faces of the tetrahedron $ABCD$ lie in a plane.

To prove this we have to see that E is the focus of a paraboloid of revolution inscribed in $ABCD$, i.e. that the quadric which is inscribed in the tetrahedron $ABCD$ and also in the cone $E(\Gamma)$ touches the plane at infinity PQR. The eight vertices of the two tetrahedra $ABCD, EPQR$, being on a twisted cubic, are such that any quadric through seven of them passes through the remaining one. It follows that the two tetrahedra are self-conjugate for the same quadric and hence, by reciprocation, that their eight faces are such that any quadric touching seven of them will touch the remaining one; the quadric inscribed in $ABCD$ and in the cone $E(\Gamma)$ touches all the faces but PQR, accordingly it touches the other face PQR and Mr Richmond's theorem is proved*.

3. I have now to make it clear that when a point E has a pedal plane for a tetrahedron $ABCD$ it is the centre of a rectangular hyperboloid S having $ABCD$ for self-conjugate tetrahedron.

The epithet rectangular implies that the linear (metrical) invariant of the quadric S is nul, i.e. that, if its equation in rectangular Cartesian coordinates is

$$ax^2 + by^2 + cz^2 + 2fyz + 2gzx + 2hxy + 2a'x + 2b'y + 2c'z + d = 0,$$

then
$$a + b + c = 0.$$

The quadric has triads of perpendicular generators.

The circle at infinity Γ being given tangentially by

$$l^2 + m^2 + n^2 = 0,$$

the quadric S is apolar to Γ, the first taken in points and the second in planes—the lineo-linear invariant of the two equations is nul†.

The section σ of S by the plane at infinity is apolar to Γ; there are triangles inscribed in σ which are self-conjugate for Γ or what is the same thing there are triangles circumscribed to Γ which are self-conjugate for σ.

* The reader will see that the reasoning is a natural extension of what might be used to establish the corresponding property of four concyclic points.

† The argument can be alternatively expressed in quadriplanar coordinates.

Digitized by Google

Writing $\Omega = 0$ as the tangential equation of the absolute, $A = 0$ as the tangential equation of the point A and so on, if P, P' are foci of a quadric of revolution inscribed in $ABCD$ (isogonal conjugates for this tetrahedron) there is an identity of the type

$$\lambda PP' + \Omega = \alpha BC + \beta CA + \gamma AB + \alpha' AD + \beta' BD + \gamma' CD,$$

the coefficients λ, α, etc. being constants.

It follows that a quadric S, apolar to the six point-pairs such as (B, C) and to Ω, is apolar to the point-pair P, P', i.e. any rectangular hyperboloid having $ABCD$ for self-conjugate tetrahedron has P, P' for conjugate points. If P has a pedal plane for $ABCD$, P' is at infinity and one of the rectangular hyperboloids having $ABCD$ for self-conjugate tetrahedron will have the plane at infinity as polar plane of P, that is, P will be its centre. Conversely, if P is the centre of such a quadric its isogonal conjugate is at infinity and it has a pedal plane for the tetrahedron $ABCD$.

4. When so much has been safely laid down, progress is easy.

Suppose for a moment that the coordinates of A, B, C, D, E are x_1, y_1, z_1, etc., then

$$A \equiv lx_1 + my_1 + nz_1 + r,$$

and we can choose the constants $\lambda, \mu, \nu, \rho, \sigma$ uniquely so that the expression

$$\lambda A^2 + \mu B^2 + \nu C^2 + \rho D^2 + \sigma E^2 \equiv \Sigma$$

contains no terms in r, i.e. so that Σ is homogeneous and of the second degree in l, m, n and $\Sigma = 0$ represents a conic at infinity. We must in fact take $\lambda, \mu, \nu, \rho, \sigma$ to be in the ratios

$$V_1 : V_2 : V_3 : V_4 : V_5,$$

where V_1 is the volume of the tetrahedron $BCDE$, etc.

Clearly any quadric whose tangential equation is of the type

$$\lambda A^2 + \mu B^2 + \nu C^2 + \rho D^2 + \sigma E^2 = 0$$

is such that the pole of the plane ABC lies on the line DE, and so on; if then the line DE cuts the plane at infinity in the point (de) and the plane ABC cuts it in the line (abc), we have the familiar figure of ten points and ten lines, each point lying on three of the lines and each line containing three of the points, viz.: the line (abc) contains the points (bc), (ca), (ab), and further the pole of the line (abc) with respect to the conic Σ is the line (de). The conic Σ is thus completely and symmetrically determined from the plane configuration, i.e. from the points A, B, C, D, E and the plane at infinity.

Digitized by Google

The quadric whose equation is

$$\lambda A^2 + \mu B^2 + \nu C^2 + \rho D^2 = 0,$$

or

$$\sigma E^2 - \Sigma = 0$$

has $ABCD$ for self-conjugate triangle, E for centre and Σ for section at infinity. If E has a pedal plane for the tetrahedron the quadric is a rectangular hyperboloid and the conic Σ taken in points is apolar to the imaginary circle at infinity Γ taken in lines; the deduction being symmetrical in statement and logically reversible, the theorem of Zeeman follows at once.

5. Let any twisted cubic through A, B, C, D, E cut the plane at infinity in the points P, Q, R; as already remarked (§ 2) the eight points mentioned are such that all quadrics through seven of them pass through the remaining one, so that there must be an identity of the type

$$\lambda A^2 + \mu B^2 + \nu C^2 + \rho D^2 + \sigma E^2 + \alpha P^2 + \beta Q^2 + \gamma R^2 = 0.$$

The equation

$$\alpha P^2 + \beta Q^2 + \gamma R^2 = 0$$

denotes a conic in the plane at infinity with PQR for a self-conjugate triangle; on account of the uniqueness of Σ we must have

$$\alpha P^2 + \beta Q^2 + \gamma R^2 \equiv -\Sigma,$$

and $\lambda, \mu, \nu, \rho, \sigma$ must have the same meanings as before.

We thus have the following theorem:

Any twisted cubic through five points A, B, C, D, E cuts the plane at infinity in three points forming a self-conjugate triangle for a fixed conic Σ, and the necessary and sufficient condition that the five points should have the Zeeman property is that there should be triangles inscribed in Σ which are self-conjugate for the imaginary circle at infinity.

The first part is of course well known.

If the triangle PQR is circumscribed to Γ there are triangles at once self-conjugate for Σ and circumscribing Γ and consequently triangles self-conjugate for Γ and inscribed in Σ. This implies that the five points A, B, C, D, E have the Zeeman property and we have Mr Richmond's theorem of § 2.

6. As regards the converse; suppose that a twisted cubic is such that any five points on it have the Zeeman property (i.e. the feet of the perpendiculars from one of them on the faces of the tetrahedron formed by the other four are coplanar), then the points P, Q, R in which the cubic cuts the plane at infinity form a triangle circumscribed to Γ.

Digitized by Google

To prove this take D, E to be fixed on the cubic and A, B, C to be variable; the conic E at infinity determined from the five points as above has the line (abc), which is arbitrary, for pole of (de), which is fixed, it has PQR for self-conjugate triangle and it is apolar to Γ in virtue of the Zeeman relation. This clearly amounts to saying that any conic having PQR for self-conjugate triangle is apolar to Γ and thence it follows immediately that Γ must be inscribed in the triangle PQR, as was to be proved.

The modifications required for higher dimensions are scarcely more than verbal.

7. The analytical condition that the five points should have the Zeeman property can be easily deduced.

Suppose that the equation of the cone joining the conic Σ to the origin is

$$ax^2 + by^2 + cz^2 + 2fyz + 2gzx + 2hxy = 0,$$

and that the five points are x_1, y_1, z_1, etc., the coordinates as usual being rectangular and Cartesian.

The tangential equation of Γ is

$$l^2 + m^2 + n^2 = 0,$$

and the Zeeman condition is

$$a + b + c = 0.$$

The tangential equation of the conic Σ or the cone above is (§ 4)

$$\lambda A^2 + \mu B^2 + \nu C^2 + \rho D^2 + \sigma E^2 = 0,$$

where $\lambda = V_1$, etc.

Writing this

$$a'l^2 + b'm^2 + c'n^2 + 2f'mn + 2g'nl + 2h'lm = 0,$$

we have

$$a' = V_1 x_1^2 + V_2 x_2^2 + V_3 x_3^2 + V_4 x_4^2 + V_5 x_5^2,$$

$$f' = V_1 y_1 z_1 + V_2 y_2 z_2 + V_3 y_3 z_3 + V_4 y_4 z_4 + V_5 y_5 z_5,$$

and $$a = b'c' - f'^2.$$

The condition being

$$a + b + c = 0,$$

we readily deduce the formula

$$\Sigma\, V_r V_s \overline{Ors}^2 = 0; \quad r, s = 1, 2, 3, 4, 5,$$

where Ors denotes the area of the triangle formed by the origin O and the points $x_r y_r z_r$, $x_s y_s z_s$.

Digitized by Google

Taking one of the points $x_r y_r z_r$, we have the equation of the pedal surface of the tetrahedron formed by the others.

The presence of the arbitrary point O in the condition need cause no surprise; it would occur in the corresponding condition in two dimensions which of course represents the fact that four points are concyclic.

8. I now pass to Mantel's theorem (§ 1), which I shall prove for general space.

Suppose that A, B, C, D, E are five points and that $\Omega = 0$ is the absolute quadric defined tangentially; the tangential equation of A is

$$lx_1 + my_1 + nz_1 + rw_1 = 0 \text{ or } A = 0,$$

where x_1, y_1, z_1, w_1 are its coordinates, here supposed homogeneous; E and E' are foci of a quadric of revolution inscribed in $ABCD$, which now means a quadric inscribed in the tangent cones from E and E' to Ω and also touching the faces of $ABCD$. The complete set A', B', C', D', E' being defined in this way, we are to prove that D' and E' are images in the plane ABC, which now means that the pole of the plane ABC for the point-pair D', E' is the same as its pole for the absolute quadric Ω.

To see this, observe that between the ten products of the type BC and Ω regarded as quadratics in l, m, n, r, there must be a relation of the type

$$\alpha BC + \beta CA + \gamma AB + \alpha' AD + \beta' BD + \gamma' CD + \lambda AE + \mu BE + \nu CE + \rho DE \equiv \Omega \quad \ldots(\text{I}),$$

the coefficients being constants.

The two quadrics

$$\alpha BC + \beta CA + \gamma AB + \alpha' AD + \beta' BD + \gamma' CD = 0$$

and

$$(\lambda A + \mu B + \nu C + \rho D) E = \Omega,$$

being the same, this quadric is inscribed in $ABCD$ and in the tangent cones from E and the point given by

$$\lambda A + \mu B + \nu C + \rho D = 0$$

to the quadric Ω. It follows now that the last point must be E', and we may write

$$E' = \lambda A + \mu B + \nu C + \rho D.$$

Similarly,

$$D' = \alpha' A + \beta' B + \gamma' C + \rho E.$$

If we write these equations in the form

$$D' = \rho E + D'',$$

$$E' = \rho D + E'',$$

Digitized by Google

the original identity becomes

$$\alpha BC + \beta CA + \gamma AB + DD'' + EE'' + \rho DE \equiv \Omega,$$

whence

$$(\rho E + D'')(\rho D + E'') + \rho(\alpha BC + \beta CA + \gamma AB) \equiv \rho\Omega + D'E'',$$

or

$$D'E' \equiv \rho\Omega + Q,$$

where Q is a homogeneous quadratic function of A, B, C.

From this form of the identity we see at once that the pole of the plane ABC for the point pair D', E' is the same as its pole for Ω; this is Mantel's theorem in its general form and the proof for any number of dimensions is exactly the same.

I now wish to make some deductions from the identity (I) in the case of Euclidean space; the coordinate w is replaced by unity and the form Ω is independent of r. On differentiation of both sides with respect to r there follows

$$\alpha(B+C) + \beta(C+A) + \gamma(A+B) + \alpha'(A+D) + \beta'(B+D) + \gamma'(C+D) + \lambda(A+E) + \mu(B+E) + \nu(C+E) + \rho(D+E) = 0,$$

or

$$(\beta+\gamma+\alpha'+\lambda)A + (\gamma+\alpha+\beta'+\mu)B + (\alpha+\beta+\gamma'+\nu)C + (\alpha'+\beta'+\gamma'+\rho)D + (\lambda+\mu+\nu+\rho)E = 0 \quad \ldots(\text{II}).$$

This being the identical relation connecting the tangential equations of A, B, C, D, E, the coefficients in the equation must be proportional to the volumes of the tetrahedra $BCDE$, etc.

Now in virtue of the equations determining the points A', B', C', D', E', viz.

$$E' \equiv \lambda A + \mu B + \nu C + \rho D,$$

$$A' \equiv \gamma B + \beta C + \alpha' D + \lambda E,$$

etc., we also have from (II) the same identity with A', B', C', D', E' in place of A, B, C, D, E respectively.

It follows that on passage from the one system of points to the other corresponding volumes remain unaltered in ratio; in other words, there being always an affine transformation (a linear one leaving the plane at infinity invariant) which changes A, B, C, D into A', B', C', D', the same will also change E into E', or the two sets can be deduced from each other by an affine transformation.

9. I shall now prove that the finite invariant planes are mutually at right angles.

It will suffice to establish this for the tetrahedra $ABCD$, $A'B'C'D'$ and I begin by observing that the perpendiculars from the vertices of the second on the faces of the first meet in a point, namely E', as follows from Mantel's theorem.

Digitized by Google

Suppose now that two tetrahedra M, N are so related to a quadric Ω that M and the polar tetrahedron N' of N with respect to Ω are in perspective; there will be a second quadric Ω' such that M and N' are polar tetrahedra with respect to Ω' and we can pass from M to N by successive reciprocations with respect to Ω' and Ω. Together these amount to a linear transformation having for invariant tetrahedron the common self-conjugate of Ω and Ω'.

Proceeding to the limit when Ω is the imaginary circle at infinity Γ, we see that passage from $ABCD$ to $A'B'C'D'$ can be effected by a linear transformation having for invariant tetrahedron a self-conjugate of Γ; this is an affine transformation of which the three finite invariant planes are mutually at right angles, and the theorem is proved.

10. The result of § 9 can be extended as follows:

If two tetrahedra are such that the perpendiculars from the vertices of the first on the faces of the second belong to a regulus, the relation is reciprocal and the affine transformation which converts one tetrahedron into the other has its three finite invariant planes mutually at right angles.

Such tetrahedra have been called orthologic and the first part of the above statement (familiar enough) is originally due, I believe, to Steiner; the second part is easily established by an argument like that used in the simpler case of § 9.

Digitized by Google

The Pedal Planes of a Tetrahedron. By Mr J. H. GRACE, Peterhouse.

[*Received* 23 May, *read* 25 July 1927.]

1. The locus of a point the feet of the perpendiculars from which on the faces of a tetrahedron are coplanar is known to be a cubic surface having nodes at the corners*: the pedal planes, by which I mean planes containing the feet of four perpendiculars, do not seem to have been much discussed. I propose to prove two theorems concerning them and make some remarks on their envelope.

What follows may be made clearer if I begin by citing some well-known results in elementary plane geometry.

The locus of a point the feet of the perpendiculars from which on the sides of a triangle lie on a straight line is the circumcircle of the triangle; there are thus two points P, Q on any line g, which possesses pedal lines for the triangle.

(i) If the pedal lines of P, Q are at right angles, g passes through a fixed point, namely the circumcentre.

(ii) In the same case if lines are drawn through P, Q parallel to their respective pedal lines the locus of their common point is the circumcircle.

(iii) The envelope of the pedal lines of a triangle is a three-cusped hypocycloid.

2. For a tetrahedron there are three points P, Q, R on any line g which have pedal planes. I shall establish the following results:

(i) If the pedal planes of P, Q, R are mutually at right angles, the line g, in general, belongs to a fixed regulus. The general case is that in which no pair of opposite edges of the tetrahedron are at right angles.

(ii) In the same case if planes are drawn through P, Q, R parallel to their respective pedal planes the locus of their common point is the curve common to the pedal locus of the tetrahedron and the quadric containing its perpendiculars.

3. I begin by recalling the fact that when a point P has a pedal plane for a tetrahedron it is the finite focus of an inscribed paraboloid of revolution; the focus at infinity is the point P' in the direction perpendicular to the pedal plane of P, for this is the direction of the axis of the paraboloid. P and P' are thus isogonal conjugates for the tetrahedron: the quadriplanar coordinates being $\alpha, \beta, \gamma, \delta$; $\alpha', \beta', \gamma', \delta'$ we have

$$\alpha\alpha' = \beta\beta' = \gamma\gamma' = \delta\delta'.$$

* See, for example, Richmond, *Camb. Phil. Soc. Proc.*, Vol. XXII. p. 35.

Digitized by Google

Now suppose that the pedal planes of three points P, Q, R on a line g are mutually at right angles and apply to g the transformation

$$\alpha, \beta, \gamma, \delta \rightarrow 1/\alpha, 1/\beta, 1/\gamma, 1/\delta;$$

the line g becomes a twisted cubic σ through the vertices A, B, C, D of the fundamental tetrahedron and the points P', Q', R' derived from P, Q, R as above, i.e., it proceeds to infinity in three directions mutually at right angles, since the pedal planes of P, Q, R have this property: it follows that all quadrics through the cubic σ are rectangular hyperboloids*.

Let E be any point on σ and let F be another point on it such that the planes ABC, DEF are perpendicular. There are four linearly independent quadrics through the six points A, B, C, D, E, F and such a tetrad may be taken to consist of three quadrics through the cubic σ and the pair of planes mentioned. All these have the rectangular hyperboloid property, and hence every quadric through the six points possess it also; in particular, the plane joining any three of the points is perpendicular to the plane joining the other three.

The plane AEF, for example, being perpendicular to BCD must contain the perpendicular from A on the face BCD, so that the line EF intersects one, and similarly all four, perpendiculars of the tetrahedron $ABCD$. The line EF must therefore be a generator of the quadric containing the perpendiculars, and, the point E being arbitrarily chosen on the cubic σ, it follows that σ must lie entirely on the quadric of the perpendiculars.

Applying the transformation

$$\alpha, \beta, \gamma, \delta \rightarrow 1/\alpha, 1/\beta, 1/\gamma, 1/\delta$$

to this result, we see that the line g is a generator of the quadric that results from the quadric of the perpendiculars under this transformation. There is no difficulty in seeing that g belongs to a definite regulus on its quadric.

To determine this regulus let planes through BC, CA, AB, perpendicular respectively to the faces, meet in D' and so on for A', B', C'. The lines AA', BB', CC', DD' belong to our regulus. In fact for the line DD' two of the points P, Q, R coincide in D, the pedal plane of which is any plane containing the perpendicular from D on the face ABC, and the other is D', the pedal plane of which is ABC: the conditions laid down for g are satisfied by the line DD'.

4. The first theorem thus established I pass to the second. Since the pedal planes of P, Q, R are mutually at right angles and

* Cf. the preceding paper § 3. The quadric has triads of perpendicular generators and its linear metrical invariant vanishes.

Digitized by Google

the point P', for example, is at infinity in the direction perpendicular to the pedal plane of P, the planes through P, Q, R parallel to their respective pedal planes are the planes $PQ'R'$, $QR'P'$, $PQ'R'$*.

These three planes meet in the point L, the locus of which is desired: L' is the point at infinity on the line PQR or g.

Consider the rectangular hyperboloids that have $ABCD$ for self-conjugate tetrahedron: one of them has its centre at P, one at Q and one at R†. I call these S_1, S_2, S_3. In regard to S_1, for example, the polar plane of P is the plane at infinity, so that the polar line of g with respect to S_1 is at infinity. Since Q, Q' and R, R', being isogonal conjugates for the tetrahedron, are conjugate pairs for the whole set of rectangular hyperboloids in hand, the polar line of g mentioned must pass through Q' and R' and so must be the line $Q'R'$. The polar plane of L' with respect to S_1 must therefore pass through Q', R' and also through P, since L' is at infinity and P is the centre of S_1; this polar plane is accordingly $PQ'R'$. It follows that L and L' are conjugate points for S_1 and similarly for S_2 and S_3; they are consequently conjugate points for all the rectangular hyperboloids of our system and isogonal conjugates for the tetrahedron $ABCD$.

When g varies the locus of L' is the section by the plane at infinity of the quadric containing the line g: applying the transformation

$$\alpha, \beta, \gamma, \delta \rightarrow 1/\alpha, 1/\beta, 1/\gamma, 1/\delta,$$

we deduce that the locus of L is the intersection of the pedal locus of the tetrahedron and the quadric containing its perpendiculars, as was to be proved.

It is clear also that L has a pedal plane which is perpendicular to the line g, since L' is at infinity on g. There is an analogous result in two dimensions.

5. So much for the general tetrahedron: the special cases, which arise when one or more pairs of opposite edges are at right angles, will now be briefly discussed. These cases present special features solely on account of the degeneration of the quadric of the perpendiculars, which can only occur when some of these lines intersect each other. If the A and D perpendiculars meet then so do the B and C and the edges AD, BC are at right angles; this is called case (α): if, in addition, the B and D perpendiculars meet it is easy to see that the four perpendiculars meet in a point and each pair of opposite edges is at right angles; this case is (β).

* The argument will be made clear by considering the analogy in two dimensions.

† Cf. § 3 of the preceding paper.

Digitized by Google

6. Case (α). One pair of opposite edges, say AD, BC are at right angles, the A and D perpendiculars meet in X and the B and C in Y. The argument of § 3, depending as it does on the line g not meeting an edge, now breaks down. It will suffice to remark that the line g instead of belonging to a regulus now belongs to one or other of two plane pencils, and that the locus of L consists of two plane sections of the pedal locus.

7. Case (β). This is more interesting. The perpendiculars of the tetrahedron now meet in a point H and the quadric of the perpendiculars is illusory, since any line through H intersects all four. The properly (and easily) modified argument of § 3 leads to the conclusion that the twisted cubic which is the transformation of g passes always through the point H, the orthocentre of the tetrahedron. The line g consequently passes through a fixed point K which is the isogonal conjugate of H; it is worth while mentioning that the feet of the perpendiculars from K on the faces of the tetrahedron are the centroids of the faces, a familiar property of this species of tetrahedron, and of course the lines AA', BB', CC', DD' of § 3 now meet in the point K. As regards the second theorem, the point L' is now anywhere at infinity and the locus of L is the pedal locus.

A further fact in this connexion is worth notice. When the perpendiculars of a tetrahedron meet in a point H, it is self-conjugate for a sphere Σ whose centre is H, and the director sphere of any inscribed quadric cuts Σ at right angles. If P is any point on the pedal locus it is the focus of an inscribed paraboloid of revolution, the director-plane of which must pass through H; the tangent at the vertex which is the pedal plane of P is one-third the distance of the director-plane from P and we readily deduce that the locus of the intersection of the pedal planes of three collinear points (in case they are mutually at right angles) is a surface homothetic with the pedal locus and of two-thirds its dimensions, the centre of similitude being the orthocentre H.

8. As regards case (β) the following alternative argument, though incomplete, may be worth notice.

I say that if the perpendiculars of a tetrahedron $ABCD$ meet in a point H the perpendiculars to the faces of their centroids meet in a point K, and any line through K cuts the pedal locus in three points, the pedal planes of which are mutually at right angles.

Suppose that four points $\alpha, \beta, \gamma, \delta$ move along four lines in space so as to describe homographic ranges on them, the same suffixes to the symbols indicating corresponding positions of the points. There are four positions of the points for which they lie in a plane π, and a tetrahedron $\alpha_r\beta_r\gamma_r\delta_r$ can be transformed into any other

Digitized by Google

$\alpha_s\beta_s\gamma_s\delta_s$ by a linear transformation having the tetrahedron formed by the planes π as invariant.

Now let g be any line and X a moving point on it, and take for $\alpha, \beta, \gamma, \delta$ the projections of X on the faces of a tetrahedron; when X varies the points describe homographic ranges on the projections of the line g on the faces and the four planes π above are the plane at infinity and the pedal planes of the three points P, Q, R in which the line g cuts the pedal locus of the tetrahedron. It follows that the pedal tetrahedron of any one point on the line g can be changed into that of any other by an affine transformation having for finite invariant planes the pedal planes of the points P, Q, R. Apply this towards the theorem enunciated above. Let $\alpha_0, \beta_0, \gamma_0, \delta_0$ be the centroids of the faces of $ABCD$, i.e. the feet of the perpendiculars from the point K, and $\alpha, \beta, \gamma, \delta$ the feet of the perpendiculars from any point X on the line g. The perpendiculars from $\alpha, \beta, \gamma, \delta$ on the faces of the tetrahedron $\alpha_0\beta_0\gamma_0\delta_0$ meet in the point X, for these faces are parallel to those of $ABCD$; referring to § 9 of the preceding paper we see that the affine transformation which converts $\alpha\beta\gamma\delta$ into $\alpha_0\beta_0\gamma_0\delta_0$ has its three finite invariant planes mutually at right angles, and these planes being, as stated, the pedal planes of P, Q, R the theorem is proved.

9. I pass now to the envelope of the pedal planes, treating only the general case.

To find the equation of the envelope suppose that π is a pedal plane and that its equation in rectangular Cartesians is

$$lx + my + nz + r = 0;$$

the planes drawn through the lines in which π cuts the faces of the tetrahedron and perpendicular to those respective faces must meet in a point, namely, that to which the pedal plane belongs. There is no difficulty in seeing that the required equation is of the form

$$U + rV = 0,$$

where U is of order four and V of order three in l, m, n. The envelope is thus a surface of class four having the plane at infinity for a tangent plane of the third order, i.e. touching it along a curve of class three given by

$$V = 0;$$

there is thus, in general, only one pedal plane in a given direction.

The surface is the reciprocal of the type of Quartic Surface (those with a triple point) classified by Rohn* and following him I regard it as fundamental to find the generators of the surface.

* Rohn, *Math. Ann.* vol. 24 (1884), p. 55.

Digitized by Google

There must be twelve at infinity, i.e. the common tangents of the curves

$$U = 0, \quad V = 0;$$

any finite generator must meet the plane at infinity where three of the above twelve co-intersect.

Now it is quite clear that the perpendiculars of the tetrahedron are all generators, the various pedal planes through them belonging to the vertices: also any plane perpendicular to an edge is a tangent plane and belongs to a point on the opposite edge, so that the six lines at infinity in planes perpendicular to the six edges of the tetrahedron are generators of the envelope. These six generators at infinity form a quadrangle and the four finite generators mentioned pass one through each corner.

Again the envelope (of class four) and the quadric of the perpendiculars have four generators in common, hence they have four others in common which belong to the opposite system on the quadric.

It may be shewn that these other finite generators are the perpendiculars of a second tetrahedron and that the remaining six generators at infinity form a second quadrangle there, through the corners of which these four new finite generators pass.

Though the two tetrahedra have much in common as regards their relations with the envelope, the same envelope does not (in general anyhow) arise from the second. It is easy to see that there are no finite generators of the envelope beyond the eight found.

The envelope of the plane through a point parallel to its pedal plane is of the same type as above.

10. No excuse has been given, and I think none is necessary, for confining the discussion to Euclidean space. The theory of the pedal plane is quite different in general projective space, the root of the matter being that in our argument the plane at infinity, as well as the absolute, plays an essential rôle.

In general space the pedal locus and the pedal envelope are reciprocals with respect to the absolute as is easily proved*, whereas in flat space the former surface is degenerate.

* Cf. Somerville, *Proc. International Congress, Cambridge*, 1912.

Digitized by Google

An Illustration of the Space Representation of Circles. By Mr J. H. GRACE, Peterhouse.

[*Received* 23 May, *read* 25 July 1927.]

1. The circles in a plane can be represented by the points of space of three dimensions in such a way that the points (point circles) of the plane correspond to the points of a quadric Q in space and two orthogonal circles in the plane correspond to two points conjugate with respect to Q*. An ∞^2 system σ of circles in the plane corresponds to a surface S in space and the reciprocal S' of S with respect to Q corresponds to a system σ' of circles in the plane closely and reciprocally related to the system σ. So far as I know only trivial examples of such systems σ and σ' have hitherto been noticed† and my object is to direct attention to an illustration drawn from the modern geometry of the triangle which is less simple and perhaps more interesting.

2. Suppose that ABC is a triangle, P any point in its plane, L, M, N the feet of the perpendiculars from P on the sides; the circle LMN is called the pedal circle of P, it is of course the auxiliary circle of the conic which has P for focus and touches the sides of the triangle and so is also the pedal circle of the second focus P' of the conic. The points P, P' are called isogonal conjugates of the triangle and if (α, β, γ), $(\alpha', \beta', \gamma')$ are their trilinear coordinates

$$\alpha\alpha' = \beta\beta' = \gamma\gamma'.$$

It is a well-known theorem (originally stated by Lemoyne) that if the point P lies on a fixed line g its pedal circle is orthogonal to a fixed circle, which I shall here call the Lemoyne circle of the line g for the triangle; my point is that the ∞^2 pedal circles of a triangle and the ∞^2 Lemoyne circles form two such reciprocal systems as σ and σ' mentioned above, i.e. they are represented in space by two surfaces S, S' reciprocal with respect to the fundamental quadric Q. The pedal circles are in fact represented by a cubic surface with four nodes, which correspond to the circles touching the sides of the triangle, and the Lemoyne circles are represented by a Steiner Quartic.

3. I begin with some general remarks relating to reciprocal (or conjugate) systems σ and σ'.

* The equation of a circle contains four coefficients and we take these or linear functions of them as coordinates of the corresponding point in space. The condition for a point circle is the vanishing of a quadratic form Q in the coefficients and two circles are orthogonal if the two sets of coefficients satisfy the conjugacy condition in regard to Q.

† The reader may refer to the treatise of Coolidge.

Digitized by Google

Let P' be any point of the surface S'; its polar plane π with respect to the fundamental quadric Q will touch the surface S, say in P: p and p' are the circles of the systems σ and σ' which correspond to the points P and P' of the surfaces S and S' that represent these systems in space. The question arises, Given the system σ in the plane what characterises the members of the system σ'?

The plane π which contains the space points conjugate to P differs from a general plane in that it touches the surface S, i.e. it cuts S in a curve having a double point (or an extra double point) and thus the set of circles of σ that are orthogonal to a member of σ' differs from the set orthogonal to a general circle in having a double element (or an extra double element), i.e. a circle that is counted twice. This is one characteristic of the system σ'.

Again the point P' of S' is conjugate not only to P but to all points of S consecutive to P, so the member p' of σ' is orthogonal not only to the element p of σ but to all consecutive elements of σ. In general the element p' is thus uniquely determined, and this is a second characteristic of the system σ'.

The first will be more immediately useful.

4. It is natural then to seek for the set of pedal circles that are orthogonal to a fixed circle.

Suppose that P, P', whose trilinear coordinates are (α, β, γ), $(1/\alpha, 1/\beta, 1/\gamma)$, are the foci of a conic Σ inscribed in the triangle ABC. The director circle of Σ cuts the self-polar circle of ABC at right angles. So if the pedal circle of P (or P') which is the auxiliary circle of Σ cuts a fixed circle at right angles the circle on PP' as diameter will cut another fixed circle at right angles and conversely. [This is a matter of elementary geometrical deduction from the fact that the square of the radius of the auxiliary circle is the arithmetic mean of the squares of the radii of the director circle and of the circle having PP', the line joining the foci, as diameter, the three circles being of course concentric.]

It follows that if the pedal circle of P and P' cuts a fixed circle at right angles the circle on PP' as diameter cuts a fixed circle at right angles, and conversely, i.e. P and P' are conjugate points for this latter circle. It suffices to set down the condition for a general conic because the points P, P' are always conjugate for the pencil of rectangular hyperbolas having ABC for self conjugate triangle and from the pencil and a general conic one circle can be compounded linearly. The condition is consequently of the type

$$p\alpha(\beta^2+\gamma^2)+q\beta(\gamma^2+\alpha^2)+r\gamma(\alpha^2+\beta^2)+2k\alpha\beta\gamma=0 \quad \ldots(\mathrm{I})$$

and represents the equation of the locus of the point P when its pedal circle is orthogonal to a fixed circle.

Digitized by Google

5. If this fixed orthogonal circle belongs to the system conjugate to the pedal circles, the set of pedal circles derived from the points on the curve defined by (I) must contain a double element and the curve (I) must have a double point. Since it is made up, however, of pairs of isogonal conjugates for the triangle ABC there will be two double points isogonally conjugate* and the cubic curve must degenerate into a line

$$l\alpha + m\beta + n\gamma = 0,$$

and a conic

$$l\beta\gamma + m\gamma\alpha + n\alpha\beta = 0,$$

the total equation being unaffected by changing α, β, γ into their reciprocals. This means that the fixed circle must be chosen so as to be orthogonal to the pedal circles of all points on a line, i.e. it must be the Lemoyne circle of that line. Thus the relation of reciprocal character between the pedal circles and the Lemoyne circles is clear, and incidentally we have proved Lemoyne's theorem, for we can choose equation (I) such that an arbitrary factor

$$l\alpha + m\beta + n\gamma$$

appears in the left-hand side.

6. The condition that two circles should be orthogonal is homogeneous and linear in the coefficients of each, and referring to equation (I) we infer that the quantities

$$X = \alpha(\beta^2 + \gamma^2), \quad Y = \beta(\gamma^2 + \alpha^2), \quad Z = \gamma(\alpha^2 + \beta^2), \quad W = 2\alpha\beta\gamma,$$

are linear functions of the coefficients of the equation of the pedal circle of the point α, β, γ. They may therefore be taken to be the coordinates of the point in space that represents this pedal circle.

The system of pedal circles *in toto* is represented by the surface whose equation is

$$W(X^2 + Y^2 + Z^2) - 2XYZ = W^3,$$

found by eliminating α, β, γ.

This cubic surface has nodes at the points

$$(1, 1, 1, 1), \quad (1, -1, -1, 1), \quad (-1, 1, -1, 1), \quad (-1, -1, 1, 1),$$

corresponding to the values of α, β, γ

$$(1, 1, 1), \quad (-1, 1, 1), \quad (1, -1, 1), \quad (1, 1, -1),$$

and thus the nodes of the surface correspond to the four circles touching the side of the triangle†.

* We may ignore the case of coincident isogonal conjugates for it must be possible to make an arbitrary pedal circle the double element. It may be noted that all curves of type (I) that pass through the incentre have a double point there and the same is true for the excentres. Cf. § 6 infra.

† Cf. preceding footnote.

Digitized by Google

For any ∞^2 set of circles it is natural to enquire as to the contained coaxal sets. They correspond to the lines lying on the representing surface. Our cubic surface contains nine lines, namely the six joining the nodes in pairs and the edges

$$BC,\ CA,\ AB,$$

of the tetrahedron of reference in our chosen coordinates X, Y, Z, W. The pedal circles of a triangle therefore include nine coaxal systems.

Six of them are the coaxal systems to which belong the various pairs of circles touching the sides. They arise from the points on the six lines bisecting the angles of the triangle. The other three (there being only three) arise from the points on the various sides of the triangle. The fact that the corresponding lines on the surface lie in a plane ($W = 0$) indicates that these three sets of coaxal circles have a common orthogonal circle; this is the circle for which the triangle is self conjugate. The organic structure of the system of pedal circles thus appears in broad outline: for closer study we should have to consider the interrelations of the four-nodal cubic surface with the fundamental quadric Q.

7. In case the fixed orthogonal circle, the existence of which leads to equation (I), is the Lemoyne circle of a line l, m, n, that equation must be identical with

$$(l\alpha + m\beta + n\gamma)\,(l\beta\gamma + m\gamma\alpha + n\alpha\beta) = 0,$$

and since the coefficients in (I) are linear functions of those of the Lemoyne circle we may take the point representing the latter circle in space to be

$$X = mn, \quad Y = nl, \quad Z = lm, \quad W = l^2 + m^2 + n^2,$$

though the coordinate system is, of course, not the same as that for the pedal circles.

The representing surface has for equation

$$Y^2Z^2 + Z^2X^2 + X^2Y^2 = XYZW$$

and accordingly it is a Steiner Quartic, as it should be, for it is the reciprocal of the four-nodal cubic above with respect to the fundamental quadric Q.

8. If the Lemoyne circle of a line l, m, n cuts a fixed circle at right angles the line envelopes a conic whose tangential equation is of the type

$$umn + vnl + wlm + t(l^2 + m^2 + n^2) = 0,$$

Digitized by Google

i.e. a conic confocal with one inscribed in the triangle ABC. When the fixed circle is a pedal circle the envelope becomes the pair of points to which the pedal circle belongs, for any line through either satisfies the conditions: Lemoyne circles orthogonal to a pedal circle thus break up into two distinct systems and, since they are represented in space as the section of a Steiner Quartic by a tangent plane, this corresponds to the well-known result that such a section consists of two conics.

9. A coaxal set of pedal circles ought, at first sight, to give rise to the conjugate set in the reciprocal or Lemoyne system, but regarded as consisting of points the Steiner Quartic (§ 7) has only three generators, viz. those given by

$$Y=0, Z=0; \quad Z=0, X=0; \quad X=0, Y=0,$$

and these correspond to the triad of lines on the cubic. The line joining two nodes on the latter surface does not yield a generator on the reciprocal because the tangent plane is the same at all points of it. The line $\beta=\gamma$, which is one of six giving rise to coaxal sets of pedal circles, occurs as a part of an infinite number of loci of type

$$p\alpha(\beta^2+\gamma^2)+q\beta(\gamma^2+\alpha^2)+r\gamma(\alpha^2+\beta^2)+2k\alpha\beta\gamma=0,$$

viz. when

$$q+r=0, p+k=0 \quad (\text{cf. } \S 6),$$

but the other part is a conic through B, C and the centres of the inscribed circle and the escribed circle opposite A. The factorisation being different, these circles orthogonal to a coaxal pedal set do not occur in the Lemoyne system as described previously.

10. A word may now be said as to the relations of the cubic surface and the Steiner Quartic with the fundamental quadric Q which represents the points of the plane.

To find the common part of the cubic surface and Q we need the pedal circles which reduce to points; the condition for a point circle being of degree two in the coordinates of the circle and these latter being of degree three in (α, β, γ), the locus of the point P in case of degeneration is of degree six and, in fact, the locus consists of the six lines joining A, B, C to I and J. The point pedal circles are the points on the lines joining the mid-points of the sides of the triangle.

It follows that the section of the cubic surface by Q consists of three conics.

The envelope of a line whose Lemoyne circle reduces to a point is of class four. Part of it is well known to be the circumcentre of the triangle, the remaining part consists of the points at infinity in the directions perpendicular to the sides. The Lemoyne

Digitized by Google

circles that arise from the lines of this envelope are the points of the nine-point circle and the points of the sides of the triangle.

Thus the section of the Steiner Quartic by Q consists of four conics.

11. Another pair of systems of circles in a plane reciprocal in the sense laid down may be alluded to.

If P, P' are the foci of an inscribed conic the coordinates of the circle on PP' as diameter may be taken to be

$$\alpha(\beta^2+\gamma^2),\quad \beta(\gamma^2+\alpha^2),\quad \gamma(\alpha^2+\beta^2),\quad 2\alpha\beta\gamma$$

(§§ 4, 6). So this system σ is represented likewise by a four-nodal cubic, but the nodes now correspond to the centres of the circles touching the sides and so are on the fundamental quadric Q.

When the point α, β, γ moves on a fixed line g the circle on PP' as diameter is orthogonal to a fixed circle; varying g we get an ∞^2 system of circles σ' reciprocal to the former and of course represented by a Steiner Quartic. This system σ' can be described more readily, in fact, than the Lemoyne system.

Let S and S' be two rectangular hyperbolas having ABC for self conjugate triangle; the pair P, P' are conjugate for all conics of the system

$$\lambda S + \mu S' + \nu g^2 = 0;$$

when P is on the line g, there is one circle included and this, being orthogonal to the circle on PP' as diameter, is a member of the system σ'. The latter system thus consists of the circles bitangent to the rectangular hyperbolas having ABC for self conjugate triangle.

It may be mentioned that the Lemoyne system consists of the circles bitangent to the parabolas with the same property.

Digitized by Google

On the generalizations of the theorems of Parseval and Riesz-Fischer. By Mr S. POLLARD, Trinity College.

[*Received* 14 June, *read* 25 July 1927.]

1. Any two real numbers p and q will be called *conjugate* if

$$\frac{1}{p}+\frac{1}{q}=1.$$

In what follows it is supposed that $p \geqslant 1$. This implies $q \geqslant 1$ and either $p=q=2$ or p and q are separated by 2.

The theorems in question are:

A. *If* $p \geqslant 2$ *and*

$$\int_{-\pi}^{\pi} |f(x)|^q \, dx < \infty, \tag{1·1}$$

then

$$\left|\frac{a_0}{\sqrt{2}}\right|^p + \sum_{n=1}^{\infty} (|a_n|^p + |b_n|^p) < \infty,$$

where, for $n = 0, 1, 2, \ldots,$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t) \cos nt \, dt, \quad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t) \sin nt \, dt,$$

and

$$\left[\left|\frac{a_0}{\sqrt{2}}\right|^p + \sum_{n=1}^{\infty} (|a_n|^p + |b_n|^p)\right]^{\frac{1}{p}} \leqslant \left[\frac{1}{\pi}\int_{-\pi}^{\pi} |f(x)|^q \, dx\right]^{\frac{1}{q}}.$$

B. *If* $p \leqslant 2$ *and*

$$\left|\frac{a_0}{\sqrt{2}}\right|^p + \sum_{n=1}^{\infty} (|a_n|^p + |b_n|^p) < \infty, \tag{1·2}$$

then there is a function $g(x)$ *such that*

$$\int_{-\pi}^{\pi} |g(x)|^q < \infty,$$

and, for $n = 0, 1, 2, \ldots,$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} g(t) \cos nt \, dt, \quad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} g(t) \sin nt \, dt. \tag{1·3}$$

Moreover

$$\left[\left|\frac{a_0}{\sqrt{2}}\right|^p + \sum_{n=1}^{\infty} (|a_n|^p + |b_n|^p)\right]^{\frac{1}{p}} \geqslant \left[\frac{1}{\pi}\int_{-\pi}^{\pi} |g(x)|^q \, dx\right]^{\frac{1}{q}}. \tag{1·4}$$

Theorem A generalizes Parseval's theorem and theorem B that of Riesz-Fischer*.

* E. W. Hobson, *Theory of Functions of a Real Variable*, 2 (1926), 599–606.

Digitized by Google

It will be observed that a distinction is drawn between the function which appears in the enunciation of B and that which appears in the enunciation of A. The two functions should not be confused.

2. As is well known these two theorems are of different types. Theorem A is of "function" type, i.e. is a direct theorem about the Fourier series of a given function; while theorem B is of "coefficient" type, i.e. is a theorem about trigonometrical series whose terms are given. It will be observed that, when the coefficients in theorem B are the Fourier coefficients of a given function $f(x)$, there is nothing in the theorem which allows us to identify this function with the function $g(x)$ given by the theorem. It is certainly true that $f(x)$ and $g(x)$ are essentially the same, but this is because of a special property of the trigonometrical orthogonal system which it does not share with all orthogonal systems.

The object of this note is to obtain a single theorem of "function" type which covers as much as possible of the ground covered by theorems A and B. It does not cover all the ground, but requires only to be supplemented by part of the ordinary Riesz-Fischer theorem, i.e. by the consequences of the usual theory of convergence on the mean.

It should be observed that the single theorem is designed in such a way that its hypothesis is as free as possible from restrictions. The only assumption made is that of integrability in the general Denjoy sense.

3. Consider:

C. *If $f(x)$ is integrable in the general Denjoy sense in $(-\pi, \pi)$ and, for $n = 0, 1, 2, \ldots$,*

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos nt\,dt, \quad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin nt\,dt,$$

then

$$\left[\frac{1}{\pi}\int_{-\pi}^{\pi} |f(x)|^q\,dx\right]^{\frac{1}{q}}$$

bears to

$$\left[\left|\frac{a_0}{\sqrt{2}}\right|^p + \sum_{n=1}^{\infty}(|a_n|^p + |b_n|^p)\right]^{\frac{1}{p}}$$

the same relation of inequality that p bears to 2 in the sense that

$$\left[\frac{1}{\pi}\int_{-\pi}^{\pi} |f(x)|^q\,dx\right]^{\frac{1}{q}} \geqslant \left[\left|\frac{a_0}{\sqrt{2}}\right|^p + \sum_{n=1}^{\infty}(|a_n|^p + |b_n|^p)\right]^{\frac{1}{q}} \quad (p \geqslant 2), \tag{3·1}$$

Digitized by Google

(3·2) $$\left[\frac{1}{\pi}\int_{-\pi}^{\pi}|f(x)|^q\,dx\right]^{\frac{1}{q}} \leqslant \left[\left|\frac{a_0}{\sqrt{2}}\right|^p + \sum_{n=1}^{\infty}(|a_n|^p + |b_n|^p)\right]^{\frac{1}{p}} \qquad (p \leqslant 2).$$

D. *If*

(3·3) $$\tfrac{1}{2}a_0^2 + \sum_{n=1}^{\infty}(a_n^2 + b_n^2) < \infty$$

then the equations

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} g(t)\cos nt\,dt, \qquad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} g(t)\sin nt\,dt \quad (n = 0, 1, 2\ldots)$$

have a solution $g(t)$ which is integrable.

Of these D is part of the ungeneralized Riesz-Fischer theorem, and C is the proposed direct theorem.

4. LEMMA. *If $f(x)$ and $g(x)$ are integrable in $(-\pi, \pi)$ in the general Denjoy sense and have the same Fourier coefficients, then* *

(4·1) $$f(x) \sim g(x) \qquad (-\pi \leqslant x \leqslant \pi).$$

Write

$$F(x) = \int_{-\pi}^{x}\{f(t) - \tfrac{1}{2}a_0\}\,dt, \qquad G(x) = \int_{-\pi}^{x}\{g(t) - \tfrac{1}{2}a_0\}\,dt,$$

so that $F(-\pi)$, $F(\pi)$, $G(-\pi)$, $G(\pi)$ all vanish and integration by parts shows that $F(x)$ and $G(x)$ have the same Fourier coefficients.

Now $F(x)$ and $G(x)$ are continuous functions and therefore, by Fejér's theorem, can be obtained from their Fourier series by the process of summation by arithmetic means. Since the Fourier series are the same it follows that $F(x)$ and $G(x)$ are the same.

Denjoy and others have, however, shown that a function which is integrable in the general Denjoy sense can be obtained almost everywhere from its indefinite integral by the process of asymptotic differentiation. Thus, provided that a function cannot have at a given point two different asymptotic differential coefficients,

$$f(x) - \tfrac{1}{2}a_0 \sim g(x) - \tfrac{1}{2}a_0 \qquad (-\pi \leqslant x \leqslant \pi),$$

which is (4·1).

To investigate the point just raised suppose that, if possible, $F(x)$ has at a point ξ two different asymptotic differential coefficients Δ_1, Δ_2. Then there are sets E_1, E_2 of unit density at ξ such that

$$\frac{F(x) - F(\xi)}{x - \xi} \to \Delta_1 \qquad (x \to \xi \text{ in } E_1),$$

$$\to \Delta_2 \qquad (x \to \xi \text{ in } E_2).$$

* By (4·1) is to be understood that $f(x) = g(x)$ throughout $-\pi \leqslant x \leqslant \pi$ except possibly at the points of a set of measure zero.

Digitized by Google

Since E_1 and E_2 are of unit density at ξ, corresponding to each positive number ϵ there is a positive number η such that

$$mE_1 I, \quad mE_2 I > (1-\epsilon)\, mI$$

for all intervals I enclosing ξ of length not greater than η. Thus

$$mE_1 E_2 I > (1-2\epsilon)\, mI,$$

which shows that $E_1 E_2$ has unit density at ξ.

This is impossible unless ξ is a limit point of $E_1 E_2$, and we now have $\dfrac{F(x)-F(\xi)}{x-\xi}$ tending to both Δ_1 and Δ_2 as $x \to \xi$ in $E_1 E_2$, which cannot be the case.

5. Suppose $p \geqslant 2$. Then unless (1·1) holds, (3·1) goes without saying, and if (1·1) holds, (3·1) is a consequence of A.

Now suppose $p \leqslant 2$. Unless (1·2) holds, (3·2) goes without saying. If (1·2) holds, then there is a function $g(x)$ such that (1·3) holds. The functions $f(x)$ and $g(x)$ have the same Fourier coefficients and, by the lemma, must coincide almost everywhere. And now (3·2) follows from (1·4).

Theorem C is thus completely established.

It will be observed that, as (1·2) with $p \leqslant 2$ implies (3·3), the special case D of B is sufficient to establish the existence of the function $g(x)$.

6. It is easily seen that C and D are in effect A and B together with the equivalence of $f(x)$ and $g(x)$. So that C and D state a little more than A and B.

In view of this a certain amount of care has to be exercised. For C and D do not extend to the case of an arbitrary orthogonal system in the same way that A and B do.

The difference is most easily illustrated in the case $p = 2$, when A and B are known to hold, with appropriate modifications, for all complete systems of normalized orthogonal functions. The same is true of D, but not of C.

For, with the aid of a construction due to Banach*, it is possible, given any integrable function $f(x)$ such that

$$\int_a^b f^2(t)\, dt = \infty, \quad \int_a^b f(t)\, dt \neq 0,$$

to obtain a complete system of normalized orthogonal functions

$$\phi_1(x), \phi_2(x), \ldots, \phi_n(x), \ldots, \tag{6·1}$$

* S. Banach, *Proc. London Math. Soc.* (2) 21 (1923), 95–97. Banach assumes that $f(x)$ is positive and so absolutely integrable. It is interesting to verify that his construction can be extended to the case of conditionally integrable functions.

Digitized by Google

with (a, b) as basic interval, such that, if

$$c_n = \int_a^b f(t)\,\phi_n(t)\,dt,$$

then, for every n,

(6·2) $$c_n = 0,$$

so that the analogue of (3·2) cannot hold.

7. Let $f(x)$ be as described and write

$$\gamma = \int_a^b f(t)\,dt,$$

so that $\gamma \neq 0$.

Take $$\psi_1(x) = \left(\frac{1}{b-a}\right)^{\frac{1}{2}},$$

$$\psi_{2r}(x) = \left(\frac{2}{b-a}\right)^{\frac{1}{2}} \cos\left(2r\pi\frac{x-a}{b-a}\right),$$

$$\psi_{2r+1}(x) = \left(\frac{2}{b-a}\right)^{\frac{1}{2}} \sin\left(2r\pi\frac{x-a}{b-a}\right) \qquad (r \geqslant 1),$$

so that the functions $\psi_r(x)$ are the trigonometrical orthogonal system adjusted to the interval (a, b) and normalized.

Since each $\psi_n(x)$ is of bounded variation, the product $f(x)\,\psi_n(x)$ is integrable. Write

$$\gamma_n = \int_a^b f(t)\,\psi_n(t)\,dt,$$

so that, by theorem C,

(7·1) $$\sum_{n=1}^{\infty} \gamma_n^2 = \int_a^b f^2(t)\,dt = \infty.$$

Take now $$\chi_n(x) = \psi_n(x) - \frac{\gamma_n}{\gamma},$$

so that

(7·2) $$\int_a^b f(t)\,\chi_n(t)\,dt = \int_a^b f(t)\,\psi_n(t)\,dt - \frac{\gamma_n}{\gamma}\int_a^b f(t)\,dt = \gamma_n - \gamma_n = 0.$$

If possible, let the functions $\chi_n(x)$ form an incomplete system, so that there is a normalized function $F(x)$ such that

$$\int_a^b F(t)\,\chi_n(t)\,dt = 0$$

for every n.

This gives

(7·3) $$\Gamma_n = \int_a^b F(t)\,\psi_n(t)\,dt = \frac{\gamma_n}{\gamma}\int_a^b F(t)\,dt.$$

Digitized by Google

Now, by theorem C,

$$\sum_{n=1}^{\infty} \Gamma_n^2 = \int_a^b F^2(t)\,dt = 1. \tag{7·4}$$

Thus, by (7·3)

$$\left\{\int_a^b F(t)\,dt\right\}^2 \sum_{n=1}^{\infty} \gamma_n^2 = \gamma^2,$$

which is impossible, in view of (7·1), as

$$\int_a^b F(t)\,dt \neq 0,$$

for otherwise, by (7·3), all the coefficients Γ_n would vanish, in defiance of (7·4).

It follows that no function such as $F(x)$ can exist and the system $\{\chi_n(t)\}$ must be complete, but not necessarily orthogonal.

The first function $\chi_1(x)$ is easily seen to vanish, since

$$\psi_1(x) - \frac{\gamma_1}{\gamma} = \left(\frac{1}{b-a}\right)^{\frac{1}{2}} - \frac{1}{\gamma}\int_a^b f(t)\left(\frac{1}{b-a}\right)^{\frac{1}{2}} dt$$

$$= \left(\frac{1}{b-a}\right)^{\frac{1}{2}} - \left(\frac{1}{b-a}\right)^{\frac{1}{2}} \frac{1}{\gamma}\gamma = 0;$$

but for all the others

$$\int_a^b \chi_n^2(t)\,dt = \int_a^b \psi_n^2(t)\,dt - \frac{2\gamma_n}{\gamma}\int_a^b \psi_n(t)\,dt + \frac{\gamma_n^2}{\gamma^2}\int_a^b 1\,dt$$

$$= 1 + (b-a)\frac{\gamma_n^2}{\gamma^2} \neq 0,$$

and the system

$$\chi_2(x),\ \chi_3(x),\ \ldots,\ \chi_n(x),\ \ldots$$

can be converted, in the usual way, into a complete normalized orthogonal system (6·1).

Since $\phi_n(x)$ is a finite linear combination of the functions $\chi_n(x)$, c_n vanishes in view of (7·2). Thus (6·2) holds, as had to be proved.

It may be observed that, since the system $\chi_n(x)$ is complete it must contain functions whose norm relative to (a, b) does not vanish and so is capable of yielding a normalized orthogonal system. The analysis given, however, shows that $\chi_1(x)$ is the only function whose norm vanishes.

Digitized by Google

Summation of the integral conjugate to the Fourier integral of finite type. By MARGARET ELEANOR GRIMSHAW, Newnham College.
(Communicated by Mr S. POLLARD.)

[*Received* 14 June, *read* 25 July 1927.]

1. The complete Denjoy-Fourier integral is of the form

$$\int_0^\infty (a_s \cos sx + b_s \sin sx)\, ds, \tag{1·1}$$

where

$$a_s = \frac{1}{\pi}\int_{-\infty}^{\infty} f(t)\cos st\,dt, \quad b_s = \frac{1}{\pi}\int_{-\infty}^{\infty} f(t)\sin st\,dt, \tag{1·2}$$

the generating function $f(t)$ being integrable in the general Denjoy sense in every finite interval, and the integrals (1·2) being interpreted as principal values or in some other suitable way.

It is possible to avoid many of the complications due to the use of the Denjoy integral by restricting $f(t)$ to some finite interval (p, q) outside which it vanishes. This is equivalent to replacing (1·1) by

$$a_s = \frac{1}{\pi}\int_p^q f(t)\cos st\,dt, \quad b_s = \frac{1}{\pi}\int_p^q f(t)\sin st\,dt. \tag{1·3}$$

When this is the case, the integral (1·1) is said to be of *finite type.*

The integral conjugate to the Denjoy-Fourier integral of finite type is

$$\int_0^\infty (b_s \cos sx - a_s \sin sx)\, ds, \tag{1·4}$$

where a_s and b_s are defined by (1·3), $f(t)$ being restricted in the way described. It is the object of this paper to discuss the Cesàro summation of the integral (1·4) at a point x in $p < x < q$. Thus we examine the behaviour, as $\omega \to \infty$, of

$$J_k(\omega) = \int_0^\omega \left(1 - \frac{s}{\omega}\right)^k (b_s \cos sx - a_s \sin sx)\, ds,$$

where k is some prescribed index of summability. We assume throughout that $k > -1$, so that

$$\int_0^\omega \left(1 - \frac{s}{\omega}\right)^k ds < \infty.$$

Digitized by Google

2. If $k > -1$, then

$$J_k(\omega) = \frac{\omega}{\pi} \int_{p-x}^{q-x} f(x+t)\, g_k(\omega t)\, dt, \tag{2·1}$$

where

$$g_k(t) = \int_0^1 (1-u)^k \sin tu\, du.$$

The method of proof is exactly similar to that used in a previous paper*, with the sine and cosine functions interchanged.

It is often convenient to write

$$J_k(\omega) = \frac{\omega}{\pi} \int_{-\infty}^{\infty} f(x+t)\, g_k(\omega t)\, dt \tag{2·2}$$

$$= \frac{\omega}{\pi} \int_0^{\infty} \chi(t)\, g_k(\omega t)\, dt,$$

where

$$\chi(t) = f(x+t) - f(x-t).$$

These forms are equivalent to (2·1) in virtue of the special property of $f(t)$. The infinite limit of integration in (2·2) can always be replaced by any finite number greater than $x-p$ and $q-x$.

3. For $k > -1$,

$$\int_0^1 (1-u)^k\, du$$

is absolutely convergent, and

$$\sin tu = \sum_{n=0}^{\infty} (-)^n \frac{(tu)^{2n+1}}{(2n+1)!},$$

the series, considered as a function of u, being uniformly convergent in (0, 1). It follows that

$$g_k(t) = \sum_{n=0}^{\infty} (-)^n \frac{t^{2n+1}}{(2n+1)!} \int_0^1 (1-u)^k\, u^{2n+1}\, du \tag{3·1}$$

$$= \sum_{n=0}^{\infty} (-)^n\, t^{2n+1} \frac{\Gamma(k+1)}{\Gamma(2n+k+3)}$$

$$= \frac{1}{t} - \frac{\Gamma(k+1)}{t^{k+1}} \sum_{n=0}^{\infty} (-)^n \frac{t^{2n+k}}{\Gamma(2n+k+1)}.$$

Thus we may write

$$g_k(t) = \frac{1}{t} - \frac{k}{t}\, \gamma_k(t), \tag{3·2}$$

where

$$\gamma_k(t) = \frac{\Gamma(k)}{t^k}\, C_k(t), \tag{3·3}$$

and

$$C_k(t) = \sum_{n=0}^{\infty} (-)^n \frac{t^{2n+k}}{\Gamma(2n+k+1)}.$$

* *Proc. Camb. Phil. Soc.*, 23 (1927), 755—767.

Digitized by Google

It may easily be shewn that if $k > 0$

$$\gamma_k(t) = \int_0^1 (1-u)^{k-1} \cos tu \, du. \tag{3·4}$$

Now it is well known that

$$\frac{d}{dt} C_{k+1}(t) = C_k(t),$$

$$|C_k(t)| \leqslant A \qquad (0 \leqslant k \leqslant 2),$$

$$\left|\frac{C_k(t)}{t^{k-2}}\right| \leqslant A \qquad (k \geqslant 2),$$

where A represents some constant, dependent upon the function under consideration. Thus

$$|\gamma_k(t)| \leqslant \frac{A}{|t|^\alpha} \qquad (k > 0), \tag{3·5}$$

where $\alpha = \min(k, 2)$.

By differentiation of (3·3),

$$\frac{1}{\Gamma(k)} \gamma_k^{(r)}(t) = \frac{C_{k-r}(t)}{t^k} - rk \frac{C_{k-r+1}(t)}{t^{k+1}} + \ldots + (-)^r k(k+1)\ldots(k+r-1) \frac{C_k(t)}{t^{k+r}},$$

where r is any positive integer not greater than k; and so

$$|\gamma_k^{(r)}(t)| \leqslant \frac{A}{|t|^\beta} \qquad (k \geqslant r), \tag{3·6}$$

where $\beta = \min(k, r+2)$.

For small values of t we may use (3·4), which gives

$$\gamma_k^{(r)}(t) = \int_0^1 (1-u)^{k-1} u^r \cos^{(r)} tu \, du,$$

whence

$$|\gamma_k^{(r)}(t)| \leqslant A. \tag{3·7}$$

Also, from (3·1), as $t \to 0$,

$$g_k(t) \sim \frac{t}{(k+1)(k+2)}. \tag{3·8}$$

4. The necessary and sufficient condition that the integral (1·4) should be summable (C, k) to s is

$$\lim_{\omega \to \infty} J_k(\omega) = s.$$

Write

$$J_k(\omega) = \frac{\omega}{\pi} \int_0^\infty \chi(t) \, g_k(\omega t) \, dt$$

$$= \frac{1}{\pi} \int_0^\infty \frac{\chi(t)}{t} \{1 - k\gamma_k(\omega t)\} \, dt,$$

by (3·2).

Digitized by Google

It would be very convenient to express $J_k(\omega)$ as the difference of the two integrals, but this is not possible in general because $\frac{\chi(t)}{t}$ is not necessarily integrable in an interval including the origin. It is indeed a simple matter to shew that if

$$\int_0^\infty \frac{\chi(t)}{t}\,dt$$

does exist then it is the sum of the conjugate integral by all Cesàro means greater than unity, and this will later appear as a particular case of a more general result. Otherwise there are two distinct methods of attacking the problem.

One method is to investigate the difference

$$J_k(\omega) - J_l(\omega), \tag{4·1}$$

where l is some other index of summability, and to find conditions under which this tends to zero. In this way we obtain two sets of points in the interval (p, q), associated respectively with Cesàro means greater than 0 and greater than 1, and such that for each point of the set there is summability by all such means or by none.

For the other method we assume that

$$\int_t^\infty \frac{\chi(u)}{u}\,du$$

is an integrable function of t in an interval including the origin. Then what follows is essentially the reduction of the condition

$$\lim_{\omega\to\infty} \frac{\omega}{\pi}\int_0^\infty \chi(t)\,g_k(\omega t)\,dt = s \tag{4·2}$$

to a neater and more manageable form.

5. We have written

$$\chi(t) = f(x+t) - f(x-t).$$

Let

$$\phi(t) = \frac{1}{\pi}\int_t^\infty \frac{\chi(u)}{u}\,du,$$

so that $\phi(t)$ vanishes for $t > x-p,\ q-x$.

By the second mean value theorem

$$\int_t^\infty \frac{\chi(u)}{u}\,du = \frac{1}{t}\int_t^\xi \chi(u)\,du \quad (t \leqslant \xi \leqslant x-p \text{ or } q-x),$$

and

$$\phi(t) = O\left(\frac{1}{t}\right) \text{ as } t \to 0. \tag{5·1}$$

Digitized by Google

Since the integrals used are to be interpreted in the general Denjoy sense, (4·2) may be replaced by the exactly equivalent condition

$$\lim_{\omega\to\infty}\lim_{\eta\to 0}\frac{\omega}{\pi}\int_\eta^\infty \chi(t)\,g_k(\omega t)\,dt = s.$$

Now

$$\omega t\,g_k(\omega t) = 1 - k\gamma_k(\omega t), \quad \frac{d}{dt}\{\omega t\,g_k(\omega t)\} = -\omega k\gamma_k'(\omega t),$$

so that

$$\int_\eta^\infty \frac{\chi(t)}{t}\,\omega t\,g_k(\omega t)\,dt = \omega\eta\pi g_k(\omega\eta)\phi(\eta) - \omega k\pi\int_\eta^\infty \phi(t)\,\gamma_k'(\omega t)\,dt.$$

By (3·8) and (5·1) the first term on the right-hand side tends to zero with η, and (4·2) is now reduced to

$$\lim_{\omega\to\infty}\lim_{\eta\to 0}\omega k\int_\eta^\infty \phi(t)\,\gamma_k'(\omega t)\,dt = -s.$$

It is easily seen that

$$\lim_{\omega\to\infty}\omega k\int_0^\infty \gamma_k'(\omega t)\,dt = -k\gamma_k(0) = -1 \qquad (k>0).$$

Thus, if $k>0$, the necessary and sufficient condition that the conjugate integral shall be summable (C, k) to s at x is

(5·2) $$\lim_{\omega\to\infty}\lim_{\eta\to 0}\omega\int_\eta^\infty \psi(t)\,\gamma_k'(\omega t)\,dt = 0,$$

where

(5·3) $$\psi(t) = \phi(t) - s \quad (0<t<x-p \text{ or } q-x),$$
$$= 0 \qquad \text{elsewhere.}$$

6. Now assume that

(6·1) $$\int_0^t \psi(u)\,du \text{ exists,}$$

and write $$\Psi_1(t) = \int_0^t \psi(u)\,du,$$

$$\Psi_2(t) = \int_0^t \Psi_1(u)\,du,$$

.....

If $$\frac{\Psi_r(t)}{t^r}\to 0 \text{ as } t\to 0$$

for some integer r, we shall say that

$$\frac{1}{\pi}\int_0^\infty \frac{\chi(u)}{u}\,du$$

Digitized by Google

exists (C, r) and has the value s. When $r=0$ this reduces to

$$\frac{1}{\pi}\int_0^\infty \frac{\chi(u)}{u}\,du = s.$$

Take $k>1$. If $X=\max(x-p, q-x)$, then the necessary and sufficient condition (5·2) for summability to s can, by (6·1), be written

$$\lim_{\omega\to\infty} \omega \int_0^X \psi(t)\,\gamma_k'(\omega t)\,dt = 0.$$

If now $0<\xi<X$,

$$\omega\int_\xi^X \psi(t)\,\gamma_k'(\omega t)\,dt = \Big[\Psi_1(t)\,\omega\gamma_k'(\omega t)\Big]_\xi^X - \omega^2\int_\xi^X \Psi_1(t)\,\gamma_k''(\omega t)\,dt.$$

Since, in the range of integration, by (3·6),

$$|\,\omega\gamma_k'(\omega t)\,| \leqslant \frac{A\omega}{(\omega\xi)^\beta},$$

where $\beta=\min(k, 3)>1$, the term between the limits tends to zero as $\omega\to\infty$.

Again, if $\xi\leqslant\lambda<\mu\leqslant X$,

$$\int_\lambda^\mu \omega^2\gamma_k''(\omega t)\,dt = \Big[\omega\gamma_k'(\omega t)\Big]_\lambda^\mu \to 0$$

as $\omega\to\infty$, by (3·6). Thus, by the convergence theorem, since $\Psi_1(t)$ is continuous,

$$\omega^2\int_\xi^X \Psi_1(t)\,\gamma_k''(\omega t)\,dt \to 0$$

as $\omega\to\infty$, which gives

$$\omega\int_\xi^X \psi(t)\,\gamma_k'(\omega t)\,dt \to 0$$

as $\omega\to\infty$, so that the condition for summability becomes, when (6·1) holds,

$$\lim_{\xi\to 0}\ \overline{\lim_{\omega\to\infty}}\left|\,\omega\int_0^\xi \psi(t)\,\gamma_k'(\omega t)\,dt\right| = 0 \qquad (k>1). \tag{6·2}$$

Repeated integration by parts yields

$$\begin{aligned}\omega\int_0^\xi \psi(t)\,\gamma_k'(\omega t)\,dt = {}& \omega\Psi_1(\xi)\,\gamma_k'(\omega\xi) - \omega^2\Psi_2(\xi)\,\gamma_k''(\omega\xi) + \dots \\ & + (-)^{r-2}\,\omega^{r-1}\,\Psi_{r-1}(\xi)\,\gamma_k^{(r-1)}(\omega\xi) \\ & + (-)^{r-1}\,\omega^r\int_0^\xi \Psi_{r-1}(t)\,\gamma_k^{(r)}(\omega t)\,dt.\end{aligned}$$

Digitized by Google

If now r is any positive integer less than $k+1$, by (3·6),

$$\omega^n \gamma_k^{(n)}(\omega\xi) \to 0 \qquad (n = 1, 2, \ldots, r-1),$$

as $\omega \to \infty$, so that (6·2) becomes

$$\lim_{\xi\to 0} \overline{\lim_{\omega\to\infty}} \left| \omega^r \int_0^\xi \Psi_{r-1}(t)\, \gamma_k^{(r)}(\omega t)\, dt \right| = 0. \tag{6·3}$$

7. A sufficient condition for summability to s is

$$\psi_n(t) = \frac{\Psi_n(t)}{t^n} \to 0 \tag{7·1}$$

as $t \to 0$, where n is the integer next below $k-1$; or, more generally,

$$\psi_r^*(t) = \frac{1}{t}\int_0^t \frac{|\Psi_{r-1}(t)|}{t^{r-1}}\, dt \to 0$$

as $t \to 0$, where r is any positive integer less than k.

We have to shew that (6·3) is satisfied. By (3·7)

$$\left| \omega^r \int_0^{1/\omega} \Psi_{r-1}(t)\, \gamma_k^{(r)}(\omega t)\, dt \right| \leqslant A\omega^r \int_0^{1/\omega} |\Psi_{r-1}(t)|\, dt$$

$$\leqslant A\psi_r^*\left(\frac{1}{\omega}\right) \to 0$$

as $\omega \to \infty$. Also

$$\left| \omega^r \int_{1/\omega}^\xi \Psi_{r-1}(t)\, \gamma_k^{(r)}(\omega t)\, dt \right| \leqslant A\omega^r \int_{1/\omega}^\xi \frac{|\Psi_{r-1}(t)|}{(\omega t)^{r+\delta}}\, dt$$

by (3·6), where $\delta = \min(k-r, 2) > 0$,

$$\leqslant \frac{A}{\omega^\delta}\int_{1/\omega}^\xi \frac{|\Psi_{r-1}(t)|}{t^{r+\delta}}\, dt$$

$$\leqslant \frac{A}{\omega^\delta}\frac{\psi_r^*(\xi)}{\xi^\delta} + \frac{A(1+\delta)M(\xi)}{\omega^\delta}\int_{1/\omega}^\xi \frac{dt}{t^{1+\delta}},$$

where $M(\xi)$ is the upper bound of $\psi_r^*(t)$ in $(0, \xi)$ and tends to zero with ξ. The first term tends to zero as $\omega \to \infty$. The second is

$$\frac{A(1+\delta)M(\xi)}{\delta}\left[1 - \frac{1}{(\omega\xi)^\delta}\right]$$

and vanishes after the operation $\lim_{\xi\to 0}\overline{\lim_{\omega\to\infty}}$.

Now (7·1) is precisely the condition that

$$\frac{1}{\pi}\int_0^\infty \frac{\chi(u)}{u}\, du \tag{7·2}$$

should have the value s (C, n). We have thus shewn that *if* (7·2) *has the value* s *in some positive Cesàro sense* (C, n), *then the conjugate integral is summable to* s, *for* $k > n+1$.

Digitized by Google

It may be noticed that if n is zero, we obtain the result that *it is sufficient for summability* (C, k), *where* $k > 1$, *that* (7·2) *should exist, and its value is the sum.*

8. An interesting sufficient condition is

$$\psi_r(t) = o(1), \quad \psi_r^*(t) = O(1) \tag{8·1}$$

as $t \to 0$, where $r < k$.

It has been shewn that for summability it is sufficient to have

$$\overline{\lim_{\omega\to\infty}} \left| \omega^r \int_0^{\xi} \Psi_{r-1}(t)\, \gamma_k^{(r)}(\omega t)\, dt \right| = 0.$$

Now if λ is any positive constant

$$\left| \omega^r \int_0^{\lambda/\omega} \Psi_{r-1}(t)\, \gamma_k^{(r)}(\omega t)\, dt \right|$$

$$\leqslant \left| \omega^r \gamma_k^{(r)}(\lambda)\, \Psi_r\left(\frac{\lambda}{\omega}\right) \right| + \left| \omega^{r+1} \int_0^{\lambda/\omega} \Psi_r(t)\, \gamma_k^{(r+1)}(\omega t)\, dt \right|$$

$$\leqslant \lambda^r A \psi_r\left(\frac{\lambda}{\omega}\right) + \lambda^r A' \omega \int_0^{\lambda/\omega} |\psi_r(t)|\, dt$$

by (3·7), and each of these terms tends to zero as $\omega \to \infty$. Thus

$$\overline{\lim_{\omega\to\infty}} \left| \omega^r \int_0^{\xi} \Psi_{r-1}(t)\, \gamma_k^{(r)}(\omega t)\, dt \right|$$

$$= \overline{\lim_{\omega\to\infty}} \left| \omega^r \int_{\lambda/\omega}^{\xi} \Psi_{r-1}(t)\, \gamma_k^{(r)}(\omega t)\, dt \right|$$

$$= \lim_{\lambda\to\infty} \overline{\lim_{\omega\to\infty}} \left| \omega^r \int_{\lambda/\omega}^{\xi} \Psi_{r-1}(t)\, \gamma_k^{(r)}(\omega t)\, dt \right|,$$

since a constant is its own limit.

As in the preceding paragraph

$$\omega^r \int_{\lambda/\omega}^{\xi} \Psi_{r-1}(t)\, \gamma_k^{(r)}(\omega t)\, dt \Big| \leqslant \frac{A}{\omega^\delta} \frac{\psi_r^*(\xi)}{\xi^\delta} + \frac{A(1+\delta) M(\xi)}{\omega^\delta} \int_{\lambda/\omega}^{\xi} \frac{dt}{t^{1-\delta}},$$

where $\delta = \min(k - r, 2)$, and $M(\xi)$ is the upper bound of $\psi_r^*(t)$ in $(0, \xi)$ and is bounded. The first term tends to zero as $\omega \to \infty$. The second is

$$\frac{A(1+\delta) M(\xi)}{\delta} \left[\frac{1}{\lambda^\delta} - \frac{1}{(\omega\xi)^\delta} \right]$$

and vanishes after the operation $\lim_{\lambda\to\infty} \overline{\lim_{\omega\to\infty}}$.

9. We now dispense with the assumption (6·1) and adopt the alternative method of attacking our problem.

Digitized by Google

We shall say that x is an *ordinary point* if

$$\chi_1(t) = \frac{1}{t}\int_0^t \chi(u)\,du = o(1), \tag{9·1}$$

and a *strictly ordinary point* if in addition

$$\chi_1^*(t) = \frac{1}{t}\int_0^t |\chi(u)|\,du = O(1) \tag{9·2}$$

as $t \to 0$.

It will be shewn that for the set of ordinary points in (p, q) the conjugate integral is summable either by all Cesàro means greater than unity or by none, while for the strictly ordinary points there is summability by all positive Cesàro means or by none. Thus we prove that, for an ordinary point and $k, l > 1$, or for a strictly ordinary point and $k, l > 0$, the difference (4·1) tends to zero.

Take $k > 0$ and let x be an ordinary point. If λ is any positive constant,

$$\omega\int_0^{\lambda/\omega} \chi(t)\,g_k(\omega t)\,dt = \lambda g_k(\lambda)\,\chi_1\left(\frac{\lambda}{\omega}\right) - \omega^2\int_0^{\lambda/\omega} t\chi_1(t)\,g_k'(\omega t)\,dt.$$

The first term tends to zero as $\omega \to \infty$. The second is in absolute value not greater than

$$A\lambda\omega\int_0^{\lambda/\omega} |\chi_1(t)|\,dt,$$

since $|g_k'(t)| \leqslant A$, and this tends to zero as $\omega \to \infty$.

We may take $\lambda = 1$, and then it appears that, as $\omega \to \infty$,

$$J_k(\omega) = \frac{\omega}{\pi}\int_{1/\omega}^{\infty} \chi(t)\,g_k(\omega t)\,dt + o(1). \tag{9·3}$$

If ξ is any positive number less than $q - x$, $x - p$, we can shew by the argument of § 6 that

$$\int_\xi^\infty \frac{\chi(t)}{t}\gamma_k(\omega t)\,dt \to 0 \tag{9·4}$$

as $\omega \to \infty$.

Let $l > 0$. We have

$$\frac{\omega}{\pi}\int_{1/\omega}^{\infty} \chi(t)\,g_k(\omega t)\,dt - \frac{\omega}{\pi}\int_{1/\omega}^{\infty} \chi(t)\,g_l(\omega t)\,dt$$
$$= \frac{1}{\pi}\int_{1/\omega}^{\infty} \frac{\chi(t)}{t}\{l\gamma_l(\omega t) - k\gamma_k(\omega t)\}\,dt,$$

and so, by (9·3) and (9·4),

$$\overline{\lim_{\omega\to\infty}}\left|J_k(\omega) - J_l(\omega)\right|$$
$$= \lim_{\xi\to 0}\overline{\lim_{\omega\to\infty}}\left|\int_{1/\omega}^{\xi} \frac{\chi(t)}{t}\{l\gamma_l(\omega t) - k\gamma_k(\omega t)\}\,dt\right|. \tag{9·5}$$

Digitized by Google

Moreover,

$$\int_{1/\omega}^{\xi} \frac{\chi(t)}{t}\gamma_k(\omega t)\,dt = \Big[\gamma_k(\omega t)\,\chi_1(t)\Big]_{1/\omega}^{\xi} - \int_{1/\omega}^{\xi} \chi_1(t)\left(\omega\gamma_k'(\omega t) - \frac{\gamma_k(\omega t)}{t}\right)dt.$$

The term between limits tends to zero as $\omega \to \infty$, by (3·5) and (9·1). Further,

$$\left|\int_{1/\omega}^{\xi} \frac{\chi_1(t)}{t}\gamma_k(\omega t)\,dt\right| \leqslant \frac{A\,M(\xi)}{\omega^{\alpha}}\int_{1/\omega}^{\xi}\frac{dt}{t^{1+\alpha}} = \frac{A\,M(\xi)}{\alpha}\left[1-\frac{1}{(\omega\xi)^{\alpha}}\right]$$

by (3·5), where $\alpha = \min(k, 2)$, $M(\xi)$ is the upper bound of $|\chi_1(t)|$ in $(0, \xi)$ and tends to zero with ξ; and the last expression vanishes after the operation $\lim_{\xi\to 0}\overline{\lim_{\omega\to\infty}}$. Finally,

$$\left|\omega\int_{1/\omega}^{\xi}\chi_1(t)\,\gamma_k'(\omega t)\,dt\right| \leqslant \frac{A\,M(\xi)}{\omega^{\beta-1}}\int_{1/\omega}^{\xi}\frac{dt}{t^{\beta}}$$

by (3·6), where $\beta = \min(k, 3)$.

Assume $k > 1$, so that $\beta > 1$. Then, exactly as above, this last expression vanishes after the operation $\lim_{\xi\to 0}\overline{\lim_{\omega\to\infty}}$. This shews that, for $k, l > 1$, at an ordinary point

$$\int_{1/\omega}^{\xi}\frac{\chi(t)}{t}\gamma_k(\omega t)\,dt, \quad \int_{1/\omega}^{\xi}\frac{\chi(t)}{t}\gamma_l(\omega t)\,dt \to 0$$

and (9·5) therefore gives

$$J_k(\omega) - J_l(\omega) \to 0. \tag{9·6}$$

Now let x be a strictly ordinary point and $k, l > 0$. We have

$$\overline{\lim_{\omega\to\infty}}\,|J_k(\omega) - J_l(\omega)| = \overline{\lim_{\omega\to\infty}}\left|\int_{\lambda/\omega}^{\xi}\frac{\chi(t)}{t}\{l\gamma_l(\omega t) - k\gamma_k(\omega t)\}\,dt\right| = \lim_{\lambda\to\infty}\overline{\lim_{\omega\to\infty}}\left|\int_{\lambda/\omega}^{\xi}\frac{\chi(t)}{t}\{l\gamma_l(\omega t) - k\gamma_k(\omega t)\}\,dt\right|.$$

By (3·5),

$$\left|\int_{\lambda/\omega}^{\xi}\frac{\chi(t)}{t}\gamma_k(\omega t)\,dt\right| \leqslant \frac{A}{\omega^{\alpha}}\int_{\lambda/\omega}^{\xi}\frac{|\chi(t)|}{t^{\alpha+1}}\,dt,$$

where $\alpha = \min(2, k) > 0$,

$$\leqslant \frac{A}{\omega^{\alpha}}\frac{\chi_1^*(\xi)}{\xi^{\alpha}} + \frac{A(\alpha+1)}{\omega^{\alpha}}\int_{\lambda/\omega}^{\xi}\frac{\chi_1^*(t)}{t^{\alpha+1}}\,dt.$$

Digitized by Google

Now, by the arguments employed in § 8, both these expressions vanish after the operation $\lim_{\lambda\to\infty}\overline{\lim_{\omega\to\infty}}$. Thus, at a strictly ordinary point, (9·6) holds for $k, l > 0$.

10. Combining these last results with those of § 7, it appears that when x is an ordinary point and

$$\frac{1}{\pi}\int_0^\infty \frac{\chi(u)}{u}\,du = s$$

in *any* positive Cesàro sense, the conjugate integral is summable to s by all Cesàro means greater than unity. If x is also a strictly ordinary point there is summability to s by all positive Cesàro means.

REFERENCES.

1. HARDY AND LITTLEWOOD, "The Allied Series of a Fourier Series." *Proc. London Math. Soc.* (2), 24 (1925), 211–246.
2. E. C. TITCHMARSH, "Conjugate Trigonometrical Integrals." *Proc. London Math. Soc.* (2), 24 (1925), 109–130.

Digitized by Google

Simplexes and other configurations upon a rational normal curve. By Mr F. P. WHITE, St John's College.

[*Received and read* 25 July 1927.]

The theorem that if two triangles be inscribed in a conic their six sides touch another conic is, of course, to be found in all the text-books; it is apparently due in the first place to Brianchon*. The further remark, that if three triangles be inscribed in a conic the three conics obtained from them in pairs have a common tangent, is to be found in Taylor's *Ancient and Modern Geometry of Conics*†; it was made independently by Wakeford.

The first generalization to the space cubic is due, so far as I have been able to discover, to von Staudt‡, who shewed that if two tetrads of points be taken on a space cubic, the eight faces are planes of a cubic developable. This theorem, again, was discovered independently by Hurwitz§ in 1875; he also pointed out that there is then on the cubic a single infinity of tetrads whose faces touch the same developable, the tetrads forming, in fact, an involution of sets of four points on the cubic curve. Meanwhile Cremona‖ had considered three n-ads on the cubic curve, and the doubly-infinite involution of sets of n points determined therefrom, and had shewn that their faces, $\binom{n}{3}$ from each n-ad, touch a surface of class $n-2$. Independently, again, Pasch¶ had considered the particular case $n=4$, in which the twelve faces of three tetrads touch a quadric.

In 1882 Emil Weyr**, who was acquainted with the work of Cremona, considered his theorem in more detail, shewing that the surface of class $n-2$ passes through $\frac{1}{2}(n-1)(n-2)$ chords of the cubic curve. For $n=5$ we get a class-cubic surface and the six chords of the cubic curve form half of a double six thereon. Weyr was, however, chiefly concerned with involutions on rational curves in the plane, and his numerous other papers†† are not much to the point.

* Brianchon, *Mémoire sur les lignes du second ordre*, Paris, 1817, p. 35 (reference taken from *Encykl. Math. Wiss.*, III, C. 1, p. 35, fn. 100).

† Cambridge, 1881, p. 360. Taylor refers to H. Picquet, *Étude géométrique des systèmes ponctuels*, Paris, 1872, but this author only gives the converse.

‡ Von Staudt, *Beitrage zur Geometrie der Lage*, 1860, p. 378.

§ Announced in *Math. Ann.*, 15, 1879, p. 14; proof *ibid.*, 20, 1882, p. 135.

‖ Cremona, *Rendiconti Lombardo* (2), 12, 1879, pp. 347–52; *Opera*, t. III, p. 441.

¶ Pasch, *Journal für Math.*, 89, 1880, p. 256.

** Weyr, *Bulletin de l'acad. roy. de Belgique* (3), 3, pp. 472–85.

†† See list in the *Royal Society Catalogue of Scientific Papers.*

Digitized by Google

Finally W. F. Meyer* gave the general theorem: The faces of k n-ads (and of the ∞^{k-1} n-ads of the involution determined by them) upon a rational normal curve of order m in space of m dimensions touch a variety of primes of dimension $k-1$ and class $\binom{n-k+1}{m-k+1}$, $(m \geqslant k \geqslant 2)$. This result is also given in Segre's article on *Mehrdimensionale Räume*†; with the extension made by F. Deruyts‡ to rational curves of order m in space of less than m dimensions we are not here concerned.

Recently Professor Kubota§ has returned to the matter, apparently without being aware of the work of Cremona, Weyr and Meyer; he has considered in particular the case of pentads upon a space cubic curve and has made a notable extension, as follows:

Taking three pentads upon the curve we get a class-cubic surface touched by the 30 faces; four pentads give four such surfaces which have the planes of a cubic developable in common; five pentads give five cubic developables which have one plane in common.

Kubota's work, like that of all writers on this matter except von Staudt, Hurwitz and Wakeford, is algebraical; and no author treats n-ads upon a rational normal curve in what would appear to be the natural manner, namely as arising from the prime sections of a curve of order n in space of n dimensions, of which the given curve is a projection. It therefore seems worth while to point out that a large number of the results hitherto obtained, including those of Kubota, with generalizations, may be very readily seen without any calculation by the help of the generation of the rational normal curve of order n from projective systems, which is explained in Veronese's paper "Princip des Projicirens und Schneidens‖."

For the sake of clearness, a few particular cases will be considered, and generalizations will be merely indicated¶.

1. *Pentads upon a cubic curve.*

Consider a rational normal quintic curve in space of five dimensions; and take any three chords l, l', l''. Then it is clear that the trisecant planes of the curve are obtained as the intersections of corresponding primes of three triply infinite projectively related systems of primes through l, l', l'' respectively.

* Meyer, *Apolarität und rationale Curven*, 1883, p. 387.

† *Encykl. Math. Wiss.*, III, C 7, p. 896. ‡ See reference in Segre, *loc. cit.*

§ Kubota, *Science Reports, Tohoku Imperial University*, 15, 1926, pp. 39–44; *Math. Zeits.*, 26, 1927, pp. 450–6.

‖ Veronese, *Math. Ann.*, 19, 1881, pp. 161–234, especially pp. 219–20.

¶ The proof of the theorem on the twelve faces of three tetrads on a space cubic curve from four dimensions is indicated in Baker's *Principles of Geometry*, vol. IV, 1925, p. 147.

Digitized by Google

Now take two arbitrary primes meeting in a solid S; a trisecant plane lying in any prime through S meets S in a line. We consider then the trisecant planes which meet S in lines.

Let P be a point of S; join it to l, l', l'' by planes ϖ_1, ϖ_2', ϖ_3'', and let ϖ_2, ϖ_3 be the planes through l corresponding to ϖ_2' and to ϖ_3'' in the projectivity mentioned above. As P varies in S we thus get three projectively related triply infinite systems of planes through l, and taking the section by an arbitrary solid Σ we get three projectivities of the space Σ. Three corresponding planes ϖ_1, ϖ_2, ϖ_3, passing through l, lie in a prime Π through l which, with its corresponding primes through l', l'', gives the trisecant plane of the quintic curve which passes through P. This trisecant plane thus projects from l on to Σ into the plane joining three corresponding points in the projectivities of Σ. If, however, the trisecant plane meets S in a line m, the solid joining it to l lies with the solids corresponding to it, in the projectively related systems through l, in the prime which projects the trisecant plane, and hence in this case the projection is a plane which contains three corresponding lines in the projectivities of Σ. This is the dual of the point of concurrence of three corresponding lines in three projectivities between the planes of a space of three dimensions and the locus of such points is known to be a sextic curve of genus three, the residual intersection of two cubic surfaces with a space cubic curve in common*.

Hence we have the result that the trisecant planes of a quintic curve in five dimensions which meet a given solid in lines project, from a chord of the curve, into the planes of a sextic developable of genus three, obtained from two class-cubic surfaces with a cubic developable in common.

On projection the quintic curve gives rise to a cubic curve in Σ, the intersections of primes through S with the quintic give an involution of sets of five points upon the cubic, determined by any two of the sets, and the ten trisecant planes lying in any such prime give the ten faces of the corresponding pentad on the cubic.

Hence the 20 *faces of any two pentads on a space cubic curve touch a sextic developable, which is touched by the faces of all the pentads of the involution which they determine.*

It is easy to see that the ∞^1 lines in which S is met by trisecant planes are the trisecant chords of the sextic curve (of genus three) in which S meets the sextic three dimensional variety of chords of the quintic.

The case of three pentads upon the cubic curve is rather simpler. These clearly arise in a similar way from the intersections

* Schur, *Math. Ann.*, 18, 1881, pp. 1–32.

Digitized by Google

of the quintic curve with three primes, which meet in a plane ϖ. The faces of the pentads of the doubly infinite involution determined by the three arise from the trisecant planes which meet ϖ, each in a point. Joining ϖ to l, l', l'' by primes Π_1, Π_2', Π_3'' and taking Π_2, Π_3, through l, corresponding to Π_2', Π_3'' respectively, we have three primes through l containing respectively the three planes ϖ_1, ϖ_2, ϖ_3 arising from any point P in ϖ. On projection we get from Π_1, Π_2, Π_3 three planes in Σ, projectively related, and the join of corresponding points in these planes is the projection of a trisecant plane of the quintic curve which meets ϖ.

Again we have the dual of a case considered by Schur and others, who obtain a cubic surface as the locus of intersections of corresponding planes of three projectively related bundles (or stars), each with a base point. Moreover, as is well known, there are six sets of corresponding planes which meet in lines, giving six lines of the cubic surface, the half of a double six. In our case we get a class-cubic surface and six lines of it; clearly these six lines arise from the chords of the quintic curve which meet ϖ, in the six intersections of ϖ with the variety of chords.

Hence we have the result:

Three pentads of points on a space cubic curve determine a doubly infinite involution of pentads, and the ten faces of each pentad touch the same class-cubic surface, which has six chords of the curve as half of a double six.

The configuration for four pentads is similarly obtained from four primes in the space of five dimensions, which have in common a line m. The threes of the four pentads determine four planes through this line and the planes common to the four class-cubic surfaces are obtained from the trisecant planes of the quintic curve which meet m. In this case we get three solids through l and sets of three planes through l, lying in them and projectively related. On projection we get, from the trisecant planes meeting m, the planes joining corresponding points of three projectively related lines, that is, the planes of a cubic developable. Hence, *the four class-cubic surfaces touching the faces of threes of four pentads upon the cubic curve have in common a cubic developable.*

Lastly, take five pentads upon the cubic curve. These arise from primes through a point P in the five-dimensional space. Through this point just one trisecant plane of the quintic can be drawn. We have thus the final result:

The five cubic developables arising from the sets of four pentads taken from the five have one plane in common.

Digitized by Google

2. *$(2r-1)$-ads upon a rational normal curve of order r.*

Precisely the same method applies to sets of $2r-1$ points upon a curve of order r in space of r dimensions. We need only state the results.

A set of $2r-1$ points in r dimensions gives, by the joins of r points, $\binom{2r-1}{r}$ primes, which we shall call the faces of the $(2r-1)$-ad.

The faces of r such $(2r-1)$-ads upon the curve of order r touch a variety of primes of dimension $r-1$ and class r, obtained from r projectively related primes, say a V^r_{r-1}.

The $r+1$ such V^r_{r-1} arising from the sets of r of $r+1$ $(2r-1)$-ads have in common a $V^{\frac{1}{2}r(r-1)}_{r-2}$.

The $r+2$ such $V^{\frac{1}{2}r(r-1)}_{r-1}$ arising from the sets of $r+1$ of $r+2$ $(2r-1)$-ads have in common a $V^{\frac{1}{3}r(r-1)(r-2)}_{r-3}$.

.....

The $r+k$ such $V^{\binom{r}{k}}_{r-k}$ arising from the sets of $r+k-1$ of $r+k$ $(2r-1)$-ads have in common a $V^{\binom{r}{k+1}}_{r-k-1}$.

.....

The $2r-1$ such $V_1{}^r$ (developables of class r) arising from the sets of $2r-2$ of $2r-1$ $(2r-1)$-ads have in common a single prime.

This general statement includes the case of § 1 for $r=3$, and for $r=2$ gives the theorem in conics mentioned in the introduction.

3. *Hexads upon a cubic curve.*

As another example, take the case of three hexads upon a space cubic curve.

These should be considered as arising from three primes and a sextic curve in space of six dimensions. The three primes intersect in a solid, which meets a plane in a prime through it in a point. We have thus to consider the trisecant planes of the sextic curve which meet a given solid S in points.

Now the trisecant planes of a rational normal sextic curve arise as the intersections of corresponding primes of four projectively related triply infinite systems of primes, the base of each being a trisecant plane of the curve. Let $\alpha, \beta, \gamma, \delta$ be the bases. Then joining a point P to $\alpha, \beta, \gamma, \delta$ by solids and proceeding as before, we get four solids through α. If these lie in a prime this prime projects from α a trisecant plane which then passes through P. Taking all points of S and projecting from α upon a solid Σ we thus get four projectivities of Σ, and the tri-

Digitized by Google

secant planes which meet S project into planes which join four corresponding and coplanar points in these projectivities. The locus of planes in Σ is thus* a surface of class 4, whose tangential equation is given by the vanishing of a determinant of four rows and columns—a surface which, following Jessop†, we may call a determinant class-quartic surface.

This surface is, however, not the most general determinant class-quartic; it contains 10 lines which are chords of the cubic curve, these arising from the 10 chords of the sextic curve which meet the solid S. (The locus of chords of the sextic is a variety of dimension 3 and of order 10.) Hence the result:

The 45 *faces of three hexads upon a space cubic curve touch a surface of class* 4 *which is the Jacobian of four quadrics, given in plane coordinates.*

It is not necessary to do more than state the results for four and for five hexads upon the cubic curve.

Taking four hexads, the four class-quartic surfaces arising from the threes thereof have in common the planes of a developable of class 6 *and of genus* 3*; taking five hexads, the five sextic developables have in common four planes*‡.

The extension to the rational normal curve of order r is at once clear. We take r sets of $2r$ points upon the curve; the $r\binom{2r}{r}$ faces touch a variety of class $r+1$; $r+1$ sets give $r+1$ such varieties, with the primes of a $V_{r-2}^{\frac{1}{2}r(r+1)}$ in common, and so on, until finally we arrive at $2r-1$ sets of $2r$ points and $r+1$ common primes.

The elementary case in which $r=2$ is interesting; the twelve sides of two quadrangles inscribed in a conic touch a curve of class 3; the three class-cubics arising from three quadrangles have three common tangents.

4. $(n+1)$*-ads upon a rational normal curve of order* n.

Another generalization of the theorem concerning tetrads upon a space cubic curve is worth special investigation, particularly as Segre§ calls attention to it in a foot-note. This is the case of simplexes upon a rational normal curve of any order.

* Schur, *loc. cit.*

† Jessop, *Quartic Surfaces*, Cambridge, 1916, Chap. IX.

‡ The points P in S through which pass trisecant planes of the sextic curve describe a symmetroid, the 10 nodes of which are the points of intersection of S with the variety of chords of the sextic. For the four hexads we get a plane in the six-dimensional space and on it a quartic curve, through the points of which pass trisecant planes; with five hexads we get a straight line and four points of it. The symmetroid, the quartic curve and the four points are the intersections of S, the plane and the line respectively with the variety of trisecant planes of the sextic, which is of dimension 5 and order 4.

§ Segre, *loc. cit.*, p. 896, fn. 375.

Digitized by Google

For this we consider the curve as the projection of a curve of order $n+1$ from a point of itself. As before, a simplex arises by projection from the intersections of this curve by a prime, and the faces of the simplex from the n-secant spaces of dimension $n-1$. Consider two such simplexes and the corresponding primes, which meet in a space S of dimension $n-1$. In a prime through S a space of dimension $n-1$ meets S in a space of dimension $n-2$. We have thus to consider the n-secant spaces of the curve of order $n+1$ which meet S in spaces of dimension $n-2$.

The n-secant spaces arise as the intersections of corresponding primes of two related ∞^n systems whose bases are respectively two points A, B of the curve. Taking any point P and joining to A, B, and taking the line through A which corresponds to the line BP through B, we have two lines through A, and any prime through the plane determined by these lines meets the corresponding prime through B in an n-secant space which passes through P.

Varying P in S and projecting from A upon a prime Σ we get two projectively related primes ϖ, ρ of Σ, corresponding points arising from the pairs of lines through A, and an n-secant space through P projects into a prime in Σ which contains two corresponding points of the primes ϖ, ρ. An n-secant space which meets S in a space of dimension $n-2$ thus gives rise in Σ to two spaces of dimension $n-2$ in ϖ, ρ which correspond in the projectivity and which lie in a prime of Σ.

The dual of all this is, first, a projectivity of the primes through two points in a space of n-dimensions, and then two corresponding lines in the projectivity which intersect; the locus of such intersections, as is well known, is a rational normal curve of order n.

Hence, in our case, *the faces of the two simplexes and of the simplexes in the involution determined by them are the osculating primes of a rational normal curve of order n.*

Three simplexes on the curve of order n arise in a similar way from three primes in the space of $n+1$ dimensions, which have in common a space of dimension $n-2$. We have to consider n-secant spaces of the curve of order $n+1$ which meet this in a space of dimension $n-3$. Projecting from A upon Σ we get two spaces α, β of dimension $n-2$, projectively related, and the faces of the simplexes arise from corresponding spaces of dimension $n-3$ in α, β which lie in a prime of Σ. We thus get a double infinity of primes of class $n-1$*.

Skipping intermediate cases, which may easily be worked out if desired, let us go on to the case of n simplexes.

* The dual is a surface of order $n-1$ arising as the locus of intersections of corresponding planes of two projectively related systems of primes through two lines.

Digitized by Google

In the space of $n+1$ dimensions we have n primes, which meet in a line l. A space of $n-1$ dimensions in a prime through l meets l in a point. We have thus to consider the n-secant spaces of the curve of order $n+1$ which pass through the points of l. We clearly get, on projection from A, two lines m, n in Σ, projectively related, two corresponding points M, N upon them arising from the same point P of l. Any prime through MN is the projection from A of an n-secant space of the curve of order $n+1$ which passes through P. The primes through MN are the tangent primes of a quadric, which is moreover degenerate tangentially into a quadric surface in the three-dimensional space of the lines m, n.

The dual theorem is perhaps easier to state:

Taking n simplexes, each formed by n osculating primes of a rational normal curve of order n, the $n(n+1)$ vertices lie upon a quadric cone of which the "vertex" is a space of dimension $n-4$.*

For $n+1$ simplexes (going back to the original form of the theorems), we have to consider the n-secant spaces through a point P, and thus on projection the primes through a line MN. Hence *the $n+1$ quadrics arising from the simplexes taken n at a time have in common all primes through a line*—the dual of the statement that quadric cones of the $(n-4)$-th kind have a generating space of dimension $n-2$ in common†.

* Cf. Segre, *loc. cit.* He does not remark that the quadric is degenerate if $n>3$. For $n=3, 2$ the cone becomes an ordinary quadric and a conic, respectively.

† The result, included in this for $n=3$, that, for four tetrads on a cubic curve, the four quadrics touching the faces of threes have a common generator, was remarked by Mr J. H. Grace.

Digitized by Google

The Specific heat of Carbon Dioxide and the form of the CO_2 *molecule.* By W. H. McCrea, B.A., Trinity College. (Communicated by Mr R. H. Fowler.)

[*Received* 28 May 1927.]

Summary. Two alternative forms of the CO_2 molecule have been suggested by various authors who have discussed the band spectrum data. The specific heat curves based on these models are considered here. It is found that neither is quite satisfactory over the whole range of temperature and we discuss the difficulties for the low temperature and high temperature portions separately. In order to get agreement for low temperatures we find it necessary to introduce a further hypothesis about the molecular model which also seems to explain one or two outstanding difficulties in interpreting the fine structure of the bands. This assumption does not make any difference at higher temperatures where we show the error in one of the curves to be of the order we should expect to be accounted for by a centrifugal stretching of the molecule.

1. The evidence of the infra-red absorption bands of CO_2 on the form of the molecule has been considered by Bjerrum*, Dennison†, Schaefer and Philipps‡, and Eucken§. All except the last of these authors agree in assigning it the form of an isosceles triangle with the carbon nucleus at the vertex, as in Fig. 1 which shows the three normal modes of vibration which will be the three fundamental frequencies of the bands. The length of the arrows gives, roughly, the relative amplitude for each particle‖.

If A, B and C are the principal moments of inertia of the system at its centre of gravity, the C-axis being parallel to OO, and the B-axis being the symmetry line, then C will be small compared with A and B, and, to a first approximation, $A\,(=B+C)$ can be taken equal to B. Thus we have the case of a symmetrical rotator whose rotational energy levels will be¶

$$W_{j,n}=\frac{h^2}{8\pi^2}\left\{j\,(j+1)\,\frac{1}{A}+n^2\left(\frac{1}{C}-\frac{1}{A}\right)\right\} \quad \ldots\ldots\ldots(1),$$

* *Verh. d. D. phys. Ges.* xvi, p. 737 (1914).
† *Phil. Mag.* i, p. 195 (1926).
‡ *Zs. f. Phys.* 36, p. 641 (1926).
§ *Zs. f. Phys.* 37, p. 714 (1926).
‖ Bjerrum (*l.c.*) obtains two solutions to the equations of motion—the present one and another—and he could not decide which should be taken.
¶ Dennison, *Nature*, Feb. 26, 1927, p. 316.

Digitized by Google

where $j = 0, 1, 2, \ldots$ $\quad n = 0, \pm 1, \pm 2, \ldots \pm j,$ or $j = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots$ $\quad n = \pm\frac{1}{2}, \pm\frac{3}{2}, \ldots \pm j,$

and, in the present case, $\frac{1}{C} \gg \frac{1}{A}$.

Using (1) the model predicts a fine structure for the bands in excellent general agreement with that observed, and gives

$$A \cong 50 \times 10^{-40} \text{ gm. cm.}^2,$$

$$C \cong 10^{-40} \text{ gm. cm.}^2,$$

assuming a gap of one line in the "n" series.

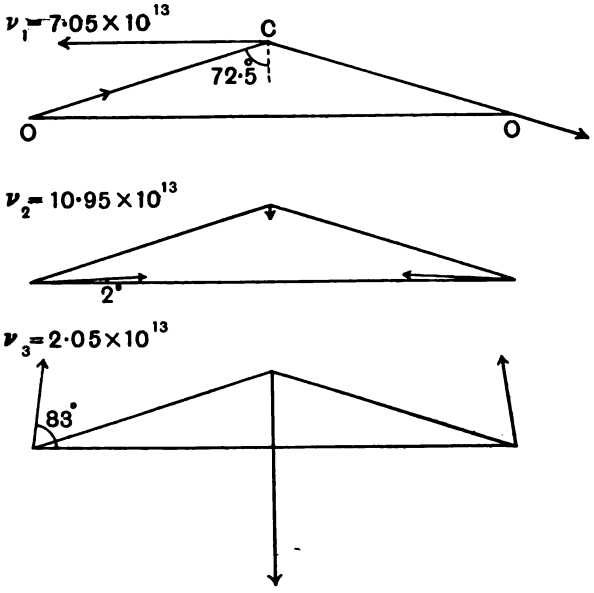


Fig. 1

Eucken, on the other hand, gives a model in which the three nuclei lie in a straight line with the carbon in the middle. His fundamental frequencies are

$\nu_1' = 6{\cdot}885 \times 10^{13}$ asymmetrical,

$\nu_2' = 3{\cdot}670 \times 10^{13}$ symmetrical,

$\nu_3' = 2{\cdot}019 \times 10^{13}$ perpendicular to the line of nuclei,

and he assigns the fine structure to a departure of the vibrations from simple harmonic type.

Digitized by Google

We cannot here go into the details of the spectral evidence which will be found in the papers quoted, for these two forms, except to say that, on the whole, it seems decidedly to favour the triangular type. It may be well, however, to mention one point noticed by Dennison, namely that, where a band is a double doublet, Eucken's explanation of the structure demands that half of it should be due to molecules excited one stage further than those molecules responsible for the other half in at least one fundamental frequency. This would predict a large difference of intensity instead of the more or less equal intensities actually observed.

2. We shall now consider the specific heat curves calculated on these two models. The former has one more rotational freedom and, giving the rotations their equipartition values, we obtain, in the usual manner, for the specific heat at constant volume,

$$\frac{C_v'}{R} = 3 + P\left(\frac{960}{T}\right) + P\left(\frac{3355}{T}\right) + P\left(\frac{5226}{T}\right) \quad \text{.....(2),}$$

$$\frac{C_v''}{R} = \frac{5}{2} + 2P\left(\frac{960}{T}\right) + P\left(\frac{1750}{T}\right) + P\left(\frac{3287}{T}\right) \quad \text{.....(3),}$$

for Dennison's and Eucken's models respectively. Here

$$P(x) = x^2 \frac{e^x}{(e^x - 1)^2},$$

and T is the absolute temperature in degrees Centigrade.

C_v' and C_v'' have been calculated for the range of temperature in which C_v has been measured and the results are given in the following table and the curves in Fig. 2.

The contribution of the ν_3 vibration to C_v' is only appreciable at high temperatures, but it is about twice the value of $C_v - C_v'$ at the end of the range.

One of Eucken's chief arguments for his model is that it does give a specific heat below $3R$ at low temperatures in agreement with experiment. We see, however, and this was not quite obvious from his paper since he was there concerned primarily with the low temperature behaviour, that (3) gives too large values at high temperatures. Now (2) and (3) both neglect the interaction of rotation and vibration, but we should expect a correction to be needed on this account, at such temperatures. This has been worked out for hydrogen* and *increased* the values calculated from a formula analogous to (2) and (3) and gave good agreement with experiment. We shall return to this below, but it is necessary to remark here that such a correction will presumably always increase the values derived from the crude formula, and, if so,

* Kemble and van Vleck, *Phys. Rev.* **xxi**, p. 653 (1923).

Digitized by Google

TABLE I.

T Absolute	C_v' Calculated from (2)	C_v'' Calculated from (3)	C_v Observed	$C_v - C_v'$
155	6·12	5·28	5·02 ± 0·4 Donath (quoted by Eucken, *l.c.*)	
197	6·33	5·72	5·83 Heuse	
273	6·75	6·70	6·67 Values given by Partington and Shilling,	
373	7·17	7·76	7·14 *The specific heats of*	
473	7·51	8·63	7·56 *gases*	
573	7·83	9·32	7·96	
673	8·10	9·86	8·34	0·24
773	8·37	10·36	8·69	0·32
873	8·64	10·73	9·03	0·39
973	8·91	11·02	9·35	0·44
1073	9·15	11·25	9·65	0·50
1173	9·38	11·46	9·92	0·54
1273	9·60	11·63	10·18	0·58
1373	9·79	11·79	10·42	0·63
1473	9·99	11·93	10·63	0·64
1573	10·17	12·05	10·83	0·66
1673	10·33	12·16	11·00	0·67
1773	10·48	12·26	11·15	0·67
1873	10·59	12·33	11·29	0·70
1973	10·69	12·38	11·40	0·71
2073	10·77	12·43	11·49	0·72
2173	10·84	12·47	11·57	0·73
2273	10·90	12·50	11·62	0·72

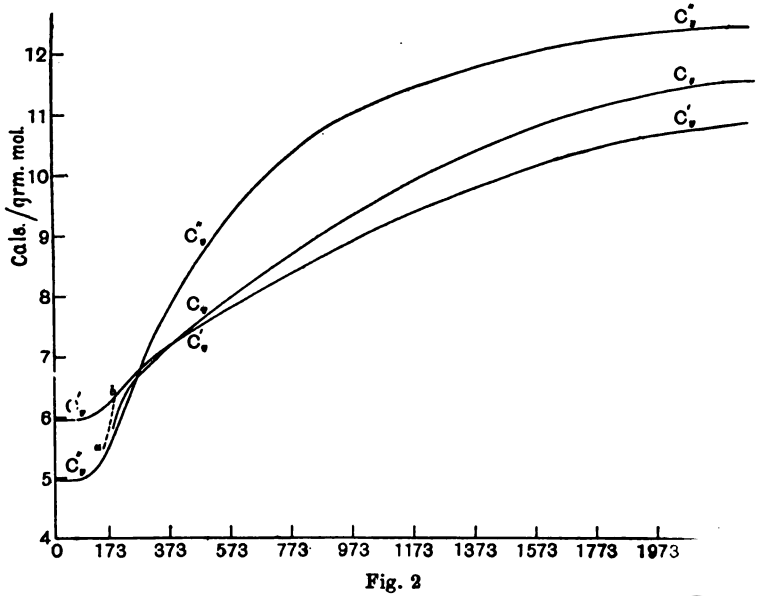


Fig. 2

Digitized by Google

any such formula predicting a value already too large must be impossible. Also, if there were a small amount of dissociation it would bring in a correction in the same direction. So it seems that Eucken's formula (3) cannot be accepted.

3. The next point is to consider whether Dennison's model will in any circumstances give values below $3R$. Now it is well known that in the case of hydrogen, on account of its small moment of inertia, the rotation does not reach its equipartition value until about 300° abs. and it was thought that here also the rotation about the small moment might show the same effect. To try this one proceeds as follows:

Equation (1) gives

$$\frac{W_{j,n}}{kT} = j(j+1)\sigma' + n^2\sigma \quad \ldots\ldots\ldots\ldots\ldots\ldots(4),$$

where $\sigma' = \dfrac{h^2}{8\pi^2 kT}\cdot\dfrac{1}{A}$, $\sigma = \dfrac{h^2}{8\pi^2 kT}\left(\dfrac{1}{C}-\dfrac{1}{A}\right)$ and $\sigma' \ll \sigma$.

The corresponding partition function is (j integral)

$$F(\sigma',\sigma) = \sum_{j=0}^{\infty}(2j+1)\,e^{-\sigma' j(j+1)} + 2\sum_{j=1}^{\infty}\sum_{n=1}^{j}(2j+1)\,e^{-\sigma' j(j+1)-\sigma n^2} \quad \ldots(5),$$

taking the usual weight factors.

We shall suppose we are dealing with such values of T that σ', but not σ, is very small (σ' will be small within a few degrees of absolute zero). It can be shown that*, for small σ' and $a \ll \dfrac{1}{\sqrt{\sigma'}}$,

$$\sum_{j=a}^{\infty}(2j+1)\,e^{-\sigma' j(j+1)} = \frac{1}{\sigma'} + O(1) \quad \ldots\ldots\ldots(6),$$

so that

$$F(\sigma',\sigma) = \frac{1}{\sigma'}\left[1 + 2\sum_{n=1}^{\infty} e^{-n^2\sigma}\right] + O(1) \ldots\ldots\ldots(7).$$

The rotational specific heat, for small σ', is therefore approximately

$$R\left\{\sigma'^2\frac{d^2}{d\sigma'^2}\log\frac{1}{\sigma'} + \sigma^2\frac{d^2}{d\sigma^2}\log\left[1+2\sum_{n=1}^{\infty}e^{-n^2\sigma}\right]\right\} \quad \ldots(8).$$

The first term is just R, i.e. the equipartition rotation about the large moments, while the second, $R\sigma^2\dfrac{d^2}{d\sigma^2}\log f(\sigma)$ (say), gives the contribution from rotation about the small moment. It is easily shown to tend to $\frac{1}{2}R$ as $\sigma \to 0$, as it must for equipartition.

* The second term on the R.H.S. is sometimes given as $O\left(\dfrac{1}{\sqrt{\sigma'}}\right)$, which is sufficient, but (6) is, in fact, true.

Digitized by Google

Values of this term have been calculated for different σ; they give a curve rising very steeply to a maximum of something like R and then approaching its asymptotic value of $\frac{1}{2}R$ from above. To fit the observed values for $T = 155°$ and 197° on to the curve would require far too small a moment of inertia and the maximum would be quite contrary to the actual facts.

In some of the attempts to get a fit for hydrogen it was found better to omit the state of no rotation; so the effect of excluding all states for which $n = 0$ was tried here, i.e. the first term (unity) was omitted from $f(\sigma)$. The resulting curve does not rise so steeply as the first. But even in this case C would have to be $0{\cdot}14 \times 10^{-40}$ gm. cm.^{2} (approx.) to get agreement and we cannot reconcile this with the values from spectral evidence, especially since, if n can never be zero, the estimate of 10^{-40} above would have to be replaced by about $1{\cdot}5 \times 10^{-40}$. Also the curve still possesses a maximum, though, indeed, it is now only about 0·56 R. Dennison (*Nature, l.c.*) has shown that half quantum numbers may be used in (1) even in the solution to the problem on the Wave Mechanics, so this was tried here. It did, indeed, give a curve rising asymptotically to $\frac{1}{2}R$ without a maximum, but required C to be about $0{\cdot}1 \times 10^{-40}$ as against 10^{-40} from the spectrum. (Here

$$F(\sigma, \sigma') = 2 \sum_{j=\frac{1}{2}}^{\infty} \sum_{n=\frac{1}{2}}^{j} 2j e^{-\sigma' j(j+1) - \sigma n^2},$$

leading finally to $$f(\sigma) = \sum_{m=0}^{\infty} e^{-m(m+1)\sigma}).$$

Lastly, in a strict application of the Wave Mechanics to the present case of a molecule which repeats itself for every half-revolution about the B-axis, the wave functions for the rotations divide into a symmetrical and an antisymmetrical series. There will be no intercombination between states belonging to different series, unless there is another wave function corresponding to some other property, say a nuclear spin, in which case the two sets of states may have different weights and we have to write*

$$\begin{aligned} F(\sigma, \sigma') &= \lambda \left\{ \sum_{j=0,2,4,\ldots}^{\infty} (2j+1)\, e^{-j(j+1)\sigma'} + \sum_{j=1,2,3,\ldots}^{\infty} \sum_{n=1}^{j} (2j+1)\, e^{-j(j+1)\sigma' - n^2\sigma} \right\} \\ &\quad + (1-\lambda) \left\{ \sum_{j=1,3,5,\ldots}^{\infty} (2j+1)\, e^{-j(j+1)\sigma'} \right. \\ &\quad \left. + \sum_{j=1,2,3,\ldots}^{\infty} \sum_{n=1}^{j} (2j+1)\, e^{-j(j+1)\sigma' - n^2\sigma} \right\} \quad \ldots(9), \end{aligned}$$

where λ has some value between 0 and 1.

* The corresponding equation for a homopolar diatomic molecule is given by Hund, *Zs. f. Phys.* 42, p. 118 (1927).

Digitized by Google

It is seen, however, that the form when σ' is very small is exactly the same as (7)* for all λ, so that there is no change from the first case considered; and this is obviously true even if there is no intercombination, for each series gives by itself the same specific heat.

For half quantum numbers there is no change in the form of $F(\sigma, \sigma')$ when this division is taken into account.

4. The failure to get agreement with experiment in this manner for any probable value of the moment C, together with the success of this particular model in other directions, led one to look for other possible ways of accounting for this behaviour at low temperatures. The following hypothesis, which seems at first sight rather arbitrary, gives reasonably good agreement for the specific heat and also seems to account for some peculiarities, remarked by Dennison† in the fine structure of some bands.

We suppose that whenever the vibration ν_3 is not excited the molecule is linear, but that whenever it is excited the molecule has the triangular form we have been considering. Having made this assumption, it is natural to suppose that when the molecule is linear there will be no rotation about the line of nuclei, and that in the other form the state of no rotation about the corresponding (C) axis is excluded.

When the A and B rotations have their equipartition values, supposing that the ν_1 and ν_2 vibrations, which in any case make negligible contributions to the specific heat below about 300° abs., are independent (see below) of the suggested change of form, the partition function for the remaining rotation and vibration is

$$1 + 2 \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} e^{-m\frac{h\nu_3}{kT} - n^2\sigma}$$

$$= 1 + \frac{2}{\left(e^{\frac{h\nu_3}{kT}} - 1\right)} \left(\sum_{n=1}^{\infty} e^{-n^2\sigma} \right) \quad \ldots\ldots\ldots\ldots\ldots(10).$$

This will, of course, give the usual values for high temperatures when the first term becomes unimportant.

Write $\dfrac{h\nu_3}{kT} = a\sigma.$

Then for $\nu_3 = 2{\cdot}05 \times 10^{13}$ and $C = 1{\cdot}01 \times 10^{-40}$,

$$a = 25,$$

* For example, $\sum_{j=0,2,4,\ldots}^{\infty} (2\gamma+1)\, e^{-j(j+1)\sigma'} = \dfrac{1}{2\sigma'} + O(1).$

† *Zs. f. Phys.* 38, p. 137 (1926).

Digitized by Google

and it is found that (10) may be written without appreciable error as

$$1 + 2e^{-a\sigma}\left(\sum_{n=1}^{\infty} e^{-n^2\sigma}\right) = \phi(\sigma) \text{ (say)} \quad \ldots\ldots\ldots(11).$$

This makes a contribution

$$R\sigma^2 \frac{d^2}{d\sigma^2} \log \phi(\sigma) \quad \ldots\ldots\ldots\ldots\ldots\ldots(12)$$

to the specific heat, which has been calculated for three temperatures, giving the results below which are indicated in Fig. 2 by the dotted curve *ab*.

TABLE II.

$T°$ Absolute	Calculated value of (12)	Resulting specific heat
158·4	0·237R	5·44
180·0	0·430R	5·82
198·0	0·628R	6·32

ab is seen to lie quite reasonably near the experimental curve (the value $T = 155°$, $C_v = 5{\cdot}02 \pm 0{\cdot}4$ is not shown in Fig. 2) and in obtaining this we have used a value of C ($1{\cdot}01 \times 10^{-40}$) of the order predicted by the optical evidence.

We may now consider briefly the optical behaviour of this molecule. Mechanically we do not expect the vibrations ν_1, ν_2 to be much affected by the contemplated change taking place in the form of the molecule, on account of the particular directions of vibration of the particles in the associated normal modes (Fig. 1) and the fact that the bending is really quite small. All we should anticipate is that ν_2 would become optically inactive in the linear form owing to the vanishing of the small vibration of the carbon nucleus.

In the ν_2 band the fine structure (Dennison, *Phil. Mag., l.c.*) is due to switches $\Delta j = 0, \pm 1$; $\Delta n = \pm 1$ in $W_{j,n}$ given by (1), i.e. lines of spacing $\left(\frac{1}{C} - \frac{1}{A}\right)\frac{h}{4\pi^2}$ and $\frac{1}{A}\frac{h}{4\pi^2}$ superposed. What is actually observed is a double doublet formed of two lines of the first spacing, each of which becomes a Bjerrum doublet of unresolved lines of the second spacing.

On the simple theory we should expect the same structure for the ν_3 band, but it is observed only as a single doublet of separation similar to that of the above Bjerrum doublets. But, on our theory,

Digitized by Google

when ν_3 is absorbed by a switch from the straight form the C rotation is not initially excited, so that n must *increase*, and thus only half of the band formerly predicted is now possible. Also, if Δn must still be unity, the only possible switch for n is $0 \rightarrow 1$ and so only *one* line of the n series will be observed. The same will be true of the combination bands $\nu_1 + \nu_3$ and $3\nu_3$ observed by Schaefer and Philipps as single doublets*.

5. It remains to consider the magnitude of the discrepancy in C_v' at high temperatures. We saw that it was probably due to a centrifugal stretching of the molecule and we appealed to the analogous phenomenon in H_2. In the diatomic case as treated by Kemble and van Vleck (*l.c.*) or Krätzer† the correction will be greater the

TABLE III.

$T°$ Absolute	Hydrogen. Specific heat (vibration + rotation)			Difference		Carbon dioxide
	(i) Observed	(ii) Calculated by Kemble and van Vleck's method	(iii) Calculated by elementary theory $R+RP\left(\frac{6982}{T}\right)$	(i)—(iii)	(ii)—(iii)	$C_v - C_v'$ from Table
673	2·12	2·08	2·01	0·11	0·07	0·24
1273	2·55	2·55	2·23	0·32	0·32	0·56
1873	3·11	3·36	2·66	0·45	0·70	0·77
2273	3·57	3·85	2·94	0·63	0·91	0·78

less the moment of inertia and the less the frequency of vibration. We notice now that in CO_2 in ν_3, the *slowest* normal mode, the particles all vibrate in directions intersecting and almost perpendicular to the C-axis, the one about which the moment is *least*. Hence this combination of vibration and rotation will, presumably, be responsible for effectively the whole correction. Further C is greater than the moment of inertia of hydrogen‡ and ν_3 is less than its fundamental frequency‡ so a C, ν_3 dipole ought to have a

* Schaefer and Philipps remark that the band which they tabulate as $2\nu_2 - \nu_1$ may alternatively be given as $\nu_2 + 2\nu_3$. It is a double doublet and forms the only possible exception to what is said above.

On our theory the two lines of the n series in ν_3, for example, will be due to switches in n of $1 \rightarrow 2$ and $2 \rightarrow 1$ (requiring again $C \cong 1{\cdot}5 \times 10^{-40}$) and we still cannot explain why more lines of the series are not observed.

† See, for example, Born, *Atommechanik*, p. 140 *et seq.*

‡ For the values of these quantities see Kemble and van Vleck.

Digitized by Google

correction of the same order as H_2. We estimate then that this combination in the triatomic molecule will give half the dipole value since the rotation has then, so to speak, only half the value. At any rate we are led to the conclusion that CO_2 should need a correction of the same order as H_2. Table III shows that this is the case and that the error $C_v - C_v'$ exhibits the right kind of variation with temperature.

No way was found for working out for a triatomic molecule a complete theory similar to Kemble and van Vleck's. It would seem that, before working out the case of a non-rigid rotator, one would have to have an exact solution for a rigid rotator with unequal moments and it will be remembered that we have hitherto neglected the small difference between A and B. The latter problem has been worked out by Witmer* but is in itself so very complicated that it seems scarcely possible to apply additional corrections.

It was thought, therefore, that it might be worth while calculating the correction for a C, ν_3 dipole. Now the formula obtained by Kemble and van Vleck contains series which are really divergent, but from the physical nature of the problem we have only to take a certain number of the terms at the beginning. This presents no difficulty for the hydrogen case where the constants involved are such that the early parts of the series behave as though they were convergent. The constants for the C, ν_3 case, however, make the terms of some of the series increase from the start so that there is no way of knowing where to stop the series. Finally, one attempted to apply to the same dipole, Krätzer's formula, with what seemed possible values of the coefficients of the small correcting terms in the energy levels. Such a large number of terms had, however, to be taken that the working evidently was not sufficiently accurate, for the result was practically identical with that of the crude theory.

Thus the qualitative comparison with known results for Hydrogen already given seems the only method left to us for estimating the correction.

6. X-ray evidence shows the molecule of solid CO_2 to be linear and, since it seems certain that the gas molecule can have the triangular form, the transition from one form to another may take place in the manner we have suggested. There may be nothing in this argument if we consider the lattice structure of solids, but it seems natural that when the lattice breaks down into the separate molecules of the gas they should leave it with the straight form.

Lastly, if both forms are possible the presence of the straight molecules would reduce the average electric moment we should

* *Proc. Nat. Acad. Sci.* 12, p. 602 (1926). The method of the old Quantum theory is used but the new theory is bound to give a similar solution.

Digitized by Google

expect supposing they all had the bent form. Experiment does indicate a non-zero electric moment though its value is not definitely settled*.

In conclusion I wish to thank Mr R. H. Fowler, for suggesting the problem and for his invaluable advice.

* Zahn, *Phys. Rev.* 27, p. 455 (1926); Weigt, *Phys. Zs.* 22, p. 643 (1921); Jona, *Phys. Zs.* 20, p. 14 (1919).

[My attention has been called by Mr F. I. G. Rawlins to some work by Kliefoth, *Zs. f. Phys.* 39, p. 402 (1926), who obtains a non-zero moment agreeing with Weigt's but by a different method. (Added in proof.)]

Digitized by Google

Eddy currents in thin circular cylinders of uniform conductivity, due to periodically changing magnetic fields, in two dimensions. By F. W. CARTER, M.A., Sc.D., St John's College. (Communicated by Mr F. P. WHITE.)

[*Received* 8 July, *read* 25 July 1927.]

1. *Introduction.*

The paper deals with the eddy currents in thin circular cylinders of uniform conducting material, due to periodic currents in conductors lying parallel to the axis of the cylinder, or to the rotation of the cylinder in a two-dimensional field of force. The first of these problems was discussed by Mr M. B. Field in a paper entitled "Eddy current losses in three-phase cable sheaths," read before the British Association at their Cambridge meeting in 1904. The solution proposed, however, although probably sufficient for the object, is mathematically defective, in that the field due to the current carried by the cable is assumed as the total field, the effect of the eddy-current field on the eddy currents themselves being left out of account.

2. *General Investigation.*

Let a be the radius of the cylindrical surface.

Let σ be the eddy-current density at any point.

Let ϕ_1 be the potential function of the eddy-current field, within the cylinder.

Let ϕ_0 be the potential function of the eddy-current field, without the cylinder.

Let ϕ' be the potential function of the disturbing field.

Let k be the product of the thickness and the specific conductivity of the cylindrical sheet.

The functions ϕ_1 and ϕ_0 satisfy Laplace's equation, and are without singularity in their respective regions. At the cylindrical surface, $r = a$,

$$\frac{\partial \phi_1}{\partial r} = \frac{\partial \phi_0}{\partial r} \quad \text{.....(1),}$$

$$4\pi\sigma = \frac{\partial \phi_1}{a\partial\theta} - \frac{\partial \phi_0}{a\partial\theta} \quad \text{.....(2),}$$

$$\frac{\partial \sigma}{a\partial\theta} = k\frac{\partial}{\partial t}\frac{\partial (\phi_1 + \phi')}{\partial r} \quad \text{.....(3).}$$

Digitized by Google

Equations 2 and 3 give

$$\frac{\partial^2 \phi_1}{\partial \theta^2} - \frac{\partial^2 \phi_0}{\partial \theta^2} = 4\pi k a^2 \frac{\partial}{\partial t} \frac{\partial (\phi_1 + \phi')}{\partial r} \quad \ldots\ldots\ldots (4).$$

Permissible forms of solution, periodic in θ and in t, and satisfying the surface equation (1), are

$$\phi_1 = \sum_n \sum_p \left(\frac{r}{a}\right)^n \{[A_n \cos pt + B_n \sin pt] \cos n\theta + [A_n' \cos pt + B_n' \sin pt] \sin n\theta\} \ldots(5),$$

$$\phi_0 = -\sum_n \sum_p \left(\frac{a}{r}\right)^n \{[A_n \cos pt + B_n \sin pt] \cos n\theta + [A_n' \cos pt + B_n' \sin pt] \sin n\theta\} \ldots(6).$$

In order to satisfy equation 4, writing $2\pi kpa = m$,

$$\sum_n \sum_p n \{[(nA_n + mB_n) \cos pt + (nB_n - mA_n) \sin pt] \cos n\theta + [(nA_n' + mB_n') \cos pt + (nB_n' - mA_n') \sin pt] \sin n\theta\}$$
$$= -2\pi k a^2 \frac{\partial^2 \phi'}{\partial t \partial r} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (7).$$

Thus if the value of $\frac{\partial^2 \phi'}{\partial t \partial r}$ at the surface of the cylinder be expanded in the form of a Fourier's series, the constants A, B, A', B', are determined by identifying the coefficients in equation (7).

3. *Disturbance due to linear alternating current.*

Let the current be $C \sin pt$, located at $(b, 0)$. Then,

$$\phi' = -2C \sin pt \sin^{-1} \frac{r \sin \theta}{(r^2 + b^2 - 2rb \cos \theta)^{\frac{1}{2}}} \quad \ldots(8),$$

$$\left(\frac{\partial^2 \phi'}{\partial t \partial r}\right)_{r=a} = 2Cp \cos pt \frac{b \sin \theta}{a^2 + b^2 - 2ab \cos \theta} \ldots\ldots\ldots\ldots (9)$$

$$= \frac{2C}{a} p \cos pt \sum_1^\infty \left(\frac{b}{a}\right)^n \sin n\theta, \text{ when } b < a,$$

$$= \frac{2C}{a} p \cos pt \sum_1^\infty \left(\frac{a}{b}\right)^n \sin n\theta, \text{ when } b > a.$$

From equation (7),

$$A = B = 0,$$

and

$$\frac{A_n'}{n} = \frac{B_n'}{m} = -\frac{2mC}{n(m^2 + n^2)} \left(\frac{b}{a}\right)^n, \text{ when } b < a,$$

$$= -\frac{2mC}{n(m^2 + n^2)} \left(\frac{a}{b}\right)^n, \text{ when } b > a.$$

Digitized by Google

Hence when $b < a$,

$$\phi_1 = -2mC \sum_1^\infty \left(\frac{br}{a^2}\right)^n \frac{n\cos pt + m\sin pt}{n(m^2+n^2)} \sin n\theta \quad \ldots(10),$$

$$\phi_0 = 2mC \sum_1^\infty \left(\frac{b}{r}\right)^n \frac{n\cos pt + m\sin pt}{n(m^2+n^2)} \sin n\theta \ldots\ldots\ldots(11),$$

and when $b > a$,

$$\phi_1 = -2mC \sum_1^\infty \left(\frac{r}{b}\right)^n \frac{n\cos pt + m\sin pt}{n(m^2+n^2)} \sin n\theta \ldots\ldots(12),$$

$$\phi_0 = 2mC \sum_1^\infty \left(\frac{a^2}{br}\right)^n \frac{n\cos pt + m\sin pt}{n(m^2+n^2)} \sin n\theta \quad \ldots\ldots(13).$$

The eddy-current density, when $b < a$, is given by (see equation (2))

$$\sigma = -\frac{mC}{\pi a} \sum_1^\infty \left(\frac{b}{a}\right)^n \frac{n\cos pt + m\sin pt}{m^2+n^2} \cos n\theta \ldots\ldots(14).$$

The mean eddy-current loss per unit length is accordingly

$$W = \int_0^1 dt \int_0^{2\pi} \frac{\sigma^2}{k} a\, d\theta$$

$$= mpC^2 \sum_1^\infty \frac{1}{m^2+n^2} \left(\frac{b}{a}\right)^{2n} \ldots\ldots\ldots\ldots(15).$$

Similarly, when $b > a$,

$$W = mpC^2 \sum_1^\infty \frac{1}{m^2+n^2} \left(\frac{a}{b}\right)^{2n} \ldots\ldots\ldots\ldots(16).$$

Note. In this case, in which the return current is not taken into account, there is a tendency to produce a current in one direction in the cylinder, which may become actual if the circuit is duly completed. A single-phase cable, having a central core, for instance, may have sheath loss if the sheath circuit is duly completed. To take account of this a term of the form

$$-2C_1\theta \sin(pt+\alpha),$$

in which $C_1 \neq C$, should be added to ϕ_0 in equations (6), (11) and (13). The corresponding loss is additive to that given by equation (15).

4. *Disturbance due to a distributed alternating current.*

It will now be shown that if the conductor carrying the alternating current is circular and such that the current density in it is a function of the distance from its centre only, the eddy current produced is the same as if the whole disturbing current were collected at the centre of the conductor.

Digitized by Google

Let $(b, 0)$ be the coordinates of the centre of the conductor, and (b', θ') those of any point on it, referred to origin at the centre of the cylinder, and let the latter of these points have coordinates (x, ϕ) referred to origin at the centre of the conductor. Let σ' be the current density at radius x in the conductor, and c the radius of the conductor. Then, from equation (10)*,

$$\phi_1 = -2m \sum_1^\infty \int_0^{2\pi} \int_0^c \sigma' x dx d\phi \left(\frac{b'r}{a^2}\right)^n \frac{n \cos pt + m \sin pt}{n(m^2+n^2)} \sin n(\theta - \theta').$$

Now $ib'^n \sin n(\theta - \theta')$ is the imaginary part of $[b'e^{i(\theta-\theta')}]^n$ or of

$$e^{ni\theta}[b + xe^{-i\phi}]^n,$$

i.e. of $$e^{ni\theta}[b^n + nb^{n-1}xe^{-i\phi} + \ldots].$$

Integrating this with respect to ϕ from 0 to 2π, gives

$$2\pi b^n e^{ni\theta},$$

of which the imaginary part is $2\pi i b^n \sin n\theta$. The same conclusion holds with reference to equations (11), (12) and (13), and the proposed result follows, since

$$C = 2\pi \int_0^c \sigma' x dx.$$

Mr Field, in the paper cited above, pointed out that in the case of stranded cable it is legitimate to assume the current density in a circular core to be a function of the distance from its centre only. For the strands are sufficiently insulated from one another to prevent sensible interchange of current with small differences of potential; and, being spiralled, the average electromotive force in the strands, due to currents in neighbouring cores, is the same for all strands at one radius. Thus all such strands are treated alike by the fields which, but for this, would tend to produce inequality in the distribution of the current.

5. *Twin core single-phase cable sheath losses.*

The field in this case is due to a current $C \sin pt$ at $(b, 0)$, and a current $-C \sin pt$ at (b, π). The combined effect is evidently to take out the even terms in equations (10) and (11), and to double the odd terms. The mean loss per unit length is therefore (cf. equation (15))

$$W = 4mpC^2 \sum_{1\,(n\,\text{odd})}^\infty \left(\frac{b}{a}\right)^{2n} \frac{1}{m^2+n^2} \quad \ldots\ldots\ldots\ldots(17),$$

which may be written

$$W = 4mpC^2 \sum_1^\infty \left(\frac{b}{a}\right)^{2n} \frac{\sin^2 \frac{n\pi}{2}}{m^2+n^2} \quad \ldots\ldots\ldots\ldots(17\,a).$$

* Strictly this formula assumes that σ' is in the same phase throughout; the restriction can however readily be shown to be unnecessary.

Digitized by Google

6. *Three core three-phase cable sheath losses.*

The field in this case is due to a current $C \sin pt$ at $(b, 0)$, a current $C \sin\left(pt - \frac{2\pi}{3}\right)$ at $\left(b, \frac{2\pi}{3}\right)$, and a current $C \sin\left(pt + \frac{2\pi}{3}\right)$ at $\left(b, -\frac{2\pi}{3}\right)$. The terms in the potential in which n is a multiple of 3 cancel; the others combine to give (cf. equation (10))

$$\phi_1 = -3mC\left\{\sum_{\substack{\nu=0 \\ n=3\nu+1}}^{\infty} \left(\frac{br}{a^2}\right)^n \frac{\sin\left(n\theta - pt + \arctan\frac{m}{n}\right)}{n\sqrt{m^2+n^2}} + \sum_{\substack{\nu=0 \\ n=3\nu+2}}^{\infty} \left(\frac{br}{a^2}\right)^n \frac{\sin\left(n\theta + pt - \arctan\frac{m}{n}\right)}{n\sqrt{m^2+n^2}}\right\} \quad \ldots\ldots(18),$$

with a corresponding expression for ϕ_0.

The mean loss per unit length of sheath is

$$W = 6mpC^2 \sum_1^{\infty} \left(\frac{b}{a}\right)^{2n} \frac{\sin^2 \frac{n\pi}{3}}{m^2+n^2} \quad \ldots\ldots(19).$$

7. *Cylinder rotating in a uniform field.*

Let p be the angular velocity, and H the field strength; then, referring to axes moving with the cylinder,

$$\phi' = -Hr\cos(\theta + pt) \quad \ldots\ldots(20),$$

$$\left(\frac{\partial^2 \phi'}{\partial t \partial r}\right)_{r=a} = Hp \sin(\theta + pt) \quad \ldots\ldots(21).$$

Hence, from equation 7, $n = 1$, and

$$-\frac{A}{m} = B = A' = \frac{B'}{m} = -\frac{maH}{1+m^2}$$

and

$$\phi_1 = -\frac{m}{\sqrt{1+m^2}} Hr \sin(\theta + pt - \arctan m) \quad \ldots\ldots(22),$$

$$\phi_0 = \frac{m}{\sqrt{1+m^2}} H\frac{a^2}{r} \sin(\theta + pt - \arctan m) \quad \ldots\ldots(23).$$

The loss per unit length is given by

$$W = \frac{H^2 a^2 p}{2} \frac{m}{1+m^2} \quad \ldots\ldots(24),$$

and this, divided by p, gives the counter-torque on the cylinder.

Digitized by Google

8. *Cylinder rotating with its axis parallel to a conductor carrying steady current C.*

In this case (cf. equation (8)),

$$\phi' = -2C \sin^{-1} \frac{r \sin(\theta + pt)}{[r^2 + b^2 - 2rb \cos(\theta + pt)]^{\frac{1}{2}}} \quad \ldots\ldots(25),$$

$$\left(\frac{\partial^2 \phi'}{\partial t \partial r}\right)_{r=a} = \frac{2C}{a} p \sum_1^\infty n \left(\frac{a}{b}\right)^n \cos n(\theta + pt) \quad \ldots\ldots(26),$$

when $b > a$.

Hence, from equation (7),

$$A_n = \frac{B_n}{m} = \frac{A_n'}{m} = -B_n = -\frac{2C}{n} \frac{m}{1+m^2} \left(\frac{a}{b}\right)^n,$$

and $$\phi_1 = -2C \frac{m}{\sqrt{1+m^2}} \sum_1^\infty \frac{1}{n} \left(\frac{r}{b}\right)^n \cos[n(\theta + pt) - \arctan m] \quad \ldots\ldots(27),$$

$$\phi_0 = 2C \frac{m}{\sqrt{1+m^2}} \sum_1^\infty \frac{1}{n} \left(\frac{a^2}{br}\right)^n \cos[n(\theta + pt) - \arctan m] \quad \ldots\ldots(28).$$

The eddy-current density is given by

$$\sigma = \frac{C}{\pi a} \frac{m}{\sqrt{1+m^2}} \sum_1^\infty \left(\frac{a}{b}\right)^n \sin[n(\theta + pt) - \arctan m] \quad \ldots\ldots(29).$$

The loss per unit length is

$$W = 2C^2 p \frac{m}{1+m^2} \sum_1^\infty \left(\frac{a}{b}\right)^{2n}$$

$$= 2C^2 p \frac{m}{1+m^2} \cdot \frac{a^2}{b^2 - a^2} \quad \ldots\ldots(30),$$

and this, divided by p, gives the counter-torque on the cylinder.

Digitized by Google

The Relation between Refractive Index and Density. By D. BURNETT, B.A., Clare College.

[*Received* 5 August, *read* 24 October 1927.]

Varying ideas of the mechanism of the propagation of light and of the constitution of matter have led at different times to different formulae for the dependence of the refractive index n of a medium on its density ρ, notably those of Laplace, $\frac{n^2-1}{\rho} = \text{constant}$, and of Lorenz* and Lorentz†, $\frac{n^2-1}{n^2+2}\frac{1}{\rho} = \text{constant}$. Recently Macdonald‡ has given a theory which seems more in accord with the modern concept of atomic structure and finds a formula $\frac{n^2-1}{n\rho} = \text{constant}$ for a non-magnetic medium. It has therefore seemed worth while to examine a number of experimental results to see whether this new formula gives as satisfactory agreement as that of Lorenz and Lorentz; it is found that it does in most cases, and is occasionally the better.

The assumptions made by the different authors about the structure of matter may be noted. Lorenz postulates a distribution of homogeneous, transparent molecules in free space, whose volumes and refractive indices do not change when the density changes; his result is that, though n and ρ are not the same for a molecule as for the matter in bulk, yet in the limiting case of infinitely long waves $\frac{n^2-1}{n^2+2}\frac{1}{\rho}$ is the same, and this relation holds approximately for the wave lengths of light. Lorentz applies the theory of electrons, with the assumption that each electron in the substance is acted on by a force, of the nature of a resistance, which is proportional to its velocity. Macdonald supposes that the difference between matter and free space can be attributed entirely to the presence of electric charges, the amplitudes of whose motion depend on the density of the matter; various simple assumptions as to the nature of this dependence lead to the formulae already quoted.

His results may be given briefly as follows: let $ie^{i(pt+\epsilon)}$ be the

* Lorenz, *Oeuvres Scientifiques*, Copenhagen, 1898, vol. I, p. 211.
† Lorentz, *Theory of Electrons*, Ch. IV.
‡ Macdonald, *Proc. Roy. Soc.* A, vol. 113, p. 237 (1926).

Digitized by Google

part of the electric current, in the direction of the electric force in the incident wave, associated with one of the charges, i_0 the value of i at the centre of a sphere, large enough to permit the assumption that it is homogeneous and has the same electric constants as the matter in bulk, yet with a radius small compared with the wave length of light; then in a non-magnetic medium $\frac{n^2-1}{n^2+2}$ is proportional to Σi_0, so that the Lorenz-Lorentz formula involves the assumption that Σi_0 is proportional to the density. It is also found that $\Sigma|i|$ is equal to (n^2-1) multiplied by a factor depending on n and on the angle of incidence of the light on the surface of the medium. In these cases *the amplitude of the incident waves is supposed unaltered* when the density varies. Macdonald then puts forward the hypothesis that the amplitude of the motion of a charge is the same in two media of the same kind but of different densities provided that *the density of electric energy is the same in both*, and this of course is not consistent with the amplitude of the incident waves being the same in the two cases. This assumption leads to $\frac{n^2-1}{n\rho}$ = constant for a non-magnetic medium and

$$\frac{n^2-1}{n\rho} = k\sqrt{1+k'^2\rho^2} + k'\sqrt{1+k^2\rho^2}$$

for a magnetic medium, k and k' being constants.

Experimental results are available for a large number of substances. In the case of gases the refractive index is so near unity that it is not always possible to differentiate between the formulae. Magri's* results for air up to pressures of 200 atmospheres are quoted by Lorentz in support of his formula, and the agreement is good. No other results have been found for such a wide range of density, but it is to be noted that recent authors† give $\frac{n-1}{\rho}$ = constant within the limit of experimental error for pressures from one to ten atmospheres; this has been found for air, oxygen, hydrogen and carbon dioxide.

In view of this uncertainty, it has seemed advisable to consider liquids rather than gases. Variation of the density of a liquid is obtained by varying the temperature, and this introduces the possibility of a shift of the absorption bands. Such an effect has been observed in specimens of glass‡ and some other crystalline

* Magri, *Phys. Zeit.* vol. 6. p. 629 (1905).

† Cf. Posejpal, *Journ. de Phys.* vol. 2, p. 87 (1921) and vol. 4, p. 451 (1923); Opladen, *Zeit. für Phys.* vol. 42, p. 160 (1927).

‡ Gibson, *Phys. Rev.* vol. 7, p. 194 (1916); Peters, *Sc. Papers of the Bureau of Standards*, 1926.

Digitized by Google

media*, and in solutions of certain colouring matters†, but information about other substances is scanty. Henri‡ has studied the absorption of most of the substances considered in this paper and he does not seem to have noted any temperature effect, so that it is probably justifiable to assume that in these cases the shift of the bands is inappreciable in the range of temperature involved.

The following table indicates the kind of results obtained: in it A and B represent the formulae $\frac{n^2-1}{n^2+2}\frac{1}{\rho}$ = constant and $\frac{n^2-1}{n\rho}$ = constant respectively.

Substance	Best Formula	Author
Water, Ice, and Steam	A	Lorenz§
Water	B	Hall and Payne‖
Alcohol (liq. and vap.)	A	Lorenz§
Alcohol (liq.)	B	Hall and Payne‖; Ketteler¶
Benzene	A or B	Landolt-Bornstein**
Ether (liq. and vap.)	A	Lorenz§
Ether (liq.)	A or B	Landolt-Bornstein**
Ethyl Bromide	A or B	Weegmann††
Aniline	B	Landolt-Bornstein**

In the cases in which an extended series of observations is available there is a steady variation in the same sense in $\frac{n^2-1}{n^2+2}\frac{1}{\rho}$ and $\frac{n^2-1}{n\rho}$. Some of the figures obtained from the measurements of Hall and Payne on pure water and alcohol may be quoted:

* Cf. Martin, *Proc. Roy. Soc.* A, vol. 96, p. 185 (1919).
† Cf. Houstoun, *Ann. der Phys.* vol. 21, p. 535 (1906).
‡ Cf. Henri, *Études de Photochimie*, Paris, 1919.
§ *Loc. cit.* p. 332.
‖ Hall and Payne, *Phys. Rev.* vol. 20, p. 249. The densities are from the *Smithsonian Tables*, 7th ed.
¶ Ketteler, *Wied. Ann.* vol. 33, p. 508 (1888).
** Landolt-Bornstein, *Tabellen.* (The values given are from different authors and only those which are consistent have been used.)
†† Weegmann, *Zeit. Ph. Chem.* vol. 2, p. 234 (1888). (A large number of similar compounds are also investigated.)

Digitized by Google

Water.

$t°$ C.	20	30	40	50	60	70
n	1·33335	1·33228	1·33087	1·32930	1·32754	1·32547
ρ	·99823	·99568	·99225	·98807	·98324	·97781
$\frac{n^2-1}{n\rho}$	·58439	·58422	·58401	·58399	·58406	·58398
$\frac{n^2-1}{n^2+2}\frac{1}{\rho}$	·20626	·20618	·20610	·20607	·20608	·20603

Alcohol.

$t°$ C.	16	20	24	28	32	36	40
$\frac{n^2-1}{n\rho}$	·7915	·7918	·7920	·7922	·7925	·7927	·7929
$\frac{n^2-1}{n^2+2}\frac{1}{\rho}$	·22402	·22412	·22420	·22427	·22434	·22442	·22450

Lorenz gives values of n and ρ for ice, water and steam; the values of the constants calculated from them are:

	Ice	Water at 10° C.	Water at 20° C.	Steam
$\frac{n^2-1}{n\rho}$	·5903	·5841	·5839	·6203
$\frac{n^2-1}{n^2+2}\frac{1}{\rho}$	·2080	·20615	·2061	·2068

It is evidently only in the case of steam that the Lorenz-Lorentz formula is the better.

The following tables were calculated from the values of n and ρ given in Landolt-Bornstein:

Aniline.

ρ	1·016	1·0216	1·0220	1·02478	1·02792
n	1·5828	1·58629	1·58632	1·58818	1·59073
$\frac{n^2-1}{n\rho}$	·9360	·9357	·9354	·9354	·9360
$\frac{n^2-1}{n^2+2}\frac{1}{\rho}$	·3288	·3286	·3285	·3285	·3286

Digitized by Google

Benzene.

ρ	·8801	·88341	·8867	·89137
n	1·49646	1·4988	1·50119	1·50381
$\frac{n^2-1}{n\rho}$	·9411	·9414	·9418	·9411
$\frac{n^2-1}{n^2+2}\frac{1}{\rho}$	·33218	·33226	·33237	·33209

Ethyl Ether.

ρ	·7141	·7157	·7166	·7183
n	1·3519	1·35032	1·35112	1·35246
$\frac{n^2-1}{n\rho}$	·8573	·8520	·8526	·8535
$\frac{n^2-1}{n^2+2}\frac{1}{\rho}$	·3028	·3008	·3011	·3015

Digitized by Google

On the Distribution of the Error of an Interpolated Value, and on the Construction of Tables. By R. A. FISHER, Sc.D., Gonville and Caius College, and Mr J. WISHART.

[*Received* 8 August, *read* 24 October 1927.]

1. *Introduction.*

Before the introduction of interpolation formulae, beyond linear interpolation by proportional parts, the presentation of the numerical values of mathematical functions was much restricted, for the labour of computation and the cost of printing, to say nothing of the inconvenience of handling a bulky volume, had to be increased quite disproportionately with every increase in accuracy. A four-figure logarithm table occupies two small pages, Chambers's seven-figure table takes 150 pages, while Vega's ten-figure table requires 300 pages twelve inches long.

The modern tendency towards compact tables used with more or less high order interpolation formulae arises from two advances: (i) the introduction of simple and adequate interpolation formulae, and (ii) the development of the theory of the remainder term. The central difference formulae of Everett not only reduce the calculation of the interpolate to a few very simple operations, of a kind suitable for machine calculations, but since they require only even differences they enable adequate differences to be presented compactly with the table. Tables of the coefficients of the second, fourth and sixth differences have been prepared by A. J. Thompson (1, 1921). Even a moderate use of differences results in a very great saving. For example in A. J. Thompson's recent twenty-figure logarithm tables (2) the values from ·9 to 1·0 occupy only 100 pages, although differences beyond the second are not required, the fourth differences being negligible. The table could in fact have been made much more compact, by the use of higher differences, although in view of the special properties of the logarithm, it is doubtful if this would be desirable. Pearson has exhibited seven-figure logarithm and anti-logarithm tables which, using second differences, together occupy only a single page (3, 1920, p. 60).

Steffensen's recent book on interpolation (4, 1927) makes available in English the author's important work on the remainder term, and so makes it possible to gauge how far high order interpolation may be carried with advantage. The contributions to the interpolate of the tabular entries and of their second, fourth, etc., differences will usually form an asymptotic series, the magnitude of the terms falling to a minimum and then increasing. This minimum should be a negligible fraction in the last decimal place

Digitized by Google

tabulated, and if so an interpolation formula will give the interpolate to the full accuracy of the table. A table should in fact always be regarded as giving not an isolated series of values of a function, but, by means of the interpolate function implicit in these, the whole continuum of its values. The possibility of adequate interpolation thus depends on the tabular intervals being sufficiently small for the remainder term to be negligible. This condition allows of the use of immensely more compact tables than the requirement that linear interpolation shall be valid, and brings with it the possibility of accurate tables of double and triple entry, without prohibitive labour either in their preparation or their use.

When the remainder term is negligible the error of the interpolate will depend only on the tabular errors, that is to say, on the differences between the exact value of the function and the corresponding value given in the table. When a table is "correct to the last figure" this error will be uniformly distributed between the limits $-0{\cdot}5$ and $+0{\cdot}5$ in the last place. Disproportionate efforts are often made to ensure this accuracy of the last place. Calculations are frequently carried to two or three figures more than are given in the final form, and the entries then cut down. Even so, it is by no means certain, owing to entries occurring that end in 50 or 500, that the final figure will in all cases be "correct"; and a lengthy series of calculations in such a doubtful case will seldom effect more improvement than to replace an error of say ·503 by one of ·497, a most trifling gain seeing that the standard error of a "correct" table is ·288. It has not been hitherto sufficiently realised that a table to a larger number of places, the last one or two of which are not necessarily correct, is capable of giving a more accurate interpolate than the same table cut down, provided that the standard error of the tabular entry is known. The expense of printing one or more figures would be more than compensated for by the increased accuracy obtainable. It is usually possible to ensure that positive and negative errors should be, in the long run, equally frequent, and that the distribution of errors shall follow approximately the normal law.

The purpose of this note is to establish the extent and nature of the errors to which the interpolate is liable, in so far as these depend only on tabular errors, in terms of those of the entries from which it is derived, whether in a table correct to the last figure or not. A statement of this error, which need not be given more than once on each page, would be a necessary and valuable addition to the many-place table that has been advocated above. It will be assumed in what follows that the standard deviations of tabular errors are all equal, within the range of the interpolation formulae, and that errors in adjacent values are independent [illegible]ther,

Digitized by Google

i.e. no correlations exist between ϵ_1 and ϵ_2, ϵ_1 and ϵ_3, etc., where ϵ denotes the error of a single value. All the formulae of interpolation by differences are included in the following general reasoning, where Lagrange's formula has been used, and the intervals assumed, for simplicity, to be equal.

2. *The Mean Variance of the Interpolate.*

Let p be the fraction of the interval (taken as 0 to 1) for which it is desired to obtain an interpolated value. Let $q = 1 - p$. Then we shall examine the contribution of any one error to the variance of the interpolate for one-, two-, three-, four-, ... point formulae and average this over the interval. The total contribution will be expressed as a polynomial in pq and we shall use the result that

$$\int_0^1 (pq)^r \, dp = \frac{(r\,!)^2}{(2r+1)\,!}.$$

1-*point Interpolation.*

The interpolate is here equal to the tabular entry, and so the ratio of the variance of the interpolate to that of the entry is unity.

2-*point Interpolation.*

We have contributions $q\epsilon$ at 0 and $p\epsilon$ at 1. The variance of the interpolate is then the average value of

$$(p^2 + q^2)\,\sigma^2 = (1 - 2pq)\,\sigma^2,$$

and our required ratio is $\frac{2}{3}$.

3-*point Interpolation.*

Contributions to the error are

$$\frac{(p-\frac{1}{2})(p-\frac{3}{2})}{2} \text{ at } -\tfrac{1}{2}, \quad (p+\tfrac{1}{2})(p-\tfrac{3}{2}) \text{ at } +\tfrac{1}{2},$$

and

$$\frac{(p+\frac{1}{2})(p-\frac{1}{2})}{2} \text{ at } +\tfrac{3}{2}.$$

We then have the total variance

$$= \left[\frac{(2p-1)^2}{64}\{(2p-3)^2 + (2p+1)^2\} + \frac{(2p+1)^2(2p-3)^2}{16}\right]\sigma^2$$

$$= \frac{\sigma^2}{32}[(1-4pq)(5-4pq) + 2(3+4pq)^2],$$

and the average value of the variance of the interpolate is

$$\frac{429}{480}\,\sigma^2,$$

the ratio of variances being ·89375.

Digitized by Google

A similar procedure gives, for 4-point interpolation, ·77566, 5-point ·89574, 6-point ·82244, 7-point ·90291, and 8-point ·84942.

3. *The Limiting Condition for High Order Interpolation.*

It is easy to show that the limit to the ratio of variances as the number of points is indefinitely increased must be unity. For the function

$$\frac{(x-a_1)(x-a_2)\ldots\ldots(x-a_{p-1})(x-a_{p+1})\ldots\ldots(x-a_n)}{(a_p-a_1)(a_p-a_2)\ldots\ldots(a_p-a_{p-1})(a_p-a_{p+1})\ldots\ldots(a_p-a_n)}$$

vanishes at all points $a_1, a_2, \ldots\ldots a_n$ except a_p, at which its value is unity. When n is large, and the intervals equal to unity, the function tends to resemble

$$\frac{\sin\pi(x-a_p)}{\pi(x-a_p)}.$$

The average contribution of any one error to the variance of the interpolate is then

$$\int_{-a_p}^{1-a_p}\frac{\sin^2\theta}{\theta^2}dx, \qquad \text{where } x=\frac{\theta}{\pi},$$

$$=\frac{1}{\pi}\int_{-\theta_p}^{\pi-\theta_p}\frac{\sin^2\theta}{\theta^2}d\theta,$$

and the average value of the total variance, expressed as a ratio to that of the tabular entry, will be

$$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin^2\theta}{\theta^2}d\theta=-\frac{1}{\pi}\left[\frac{\sin^2\theta}{\theta}\right]_{-\infty}^{\infty}+\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{2\sin\theta\cos\theta}{\theta}d\theta,$$

of which the first term vanishes, while the second is

$$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin\phi}{\phi}d\phi=1. \qquad \ldots\ldots(3{\cdot}1)$$

We have, then, the result that the average standard error of the interpolate is always somewhat less than that of the tabular entry, being least for linear interpolation, and the advantage of a system of tabulation which is not concerned with the accuracy of the last figure recorded, but which states the standard error of entry, is apparent.

For certain functions it is possible to tabulate better values for the even differences than those derived from the tabular values. If exact values for the differences are employed in the interpolation formula, errors will only be introduced by the part calculated from the two adjacent tabular values, and the distribution of

Digitized by Google

errors will be exactly the same as for 2-point interpolation. This condition is nearly realised when even differences are given correct to the last digit of the tabular values.

4. *The Error Distribution for High Order Interpolation.*

When we turn to examine not merely the variance but the form of the distribution of the error in the interpolate, a normally distributed tabular error is seen to have an added advantage beyond that due to the increased accuracy attainable when the table is not cut down to be "correct to the last place." In what follows n, the number of points, will be assumed to be large, and the theory put forward may be regarded as that to which we tend as the number of differences used in interpolation is increased. It will first be necessary to find the average values of the powers of the several errors. For the average value of the 4th powers of the tabular errors we require

$$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin^4\theta}{\theta^4}\,d\theta=\frac{1}{3!\,\pi}\int_{-\infty}^{\infty}(2^3-4)\frac{\sin\theta}{\theta}\,d\theta=\frac{2}{3}. \quad \ldots(4{\cdot}1)$$

Similarly $$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin^6\theta}{\theta^6}\,d\theta=\frac{1}{5!}(3^5-6.2^5+15)=\frac{11}{20}, \quad \ldots\ldots(4{\cdot}2)$$

while for the eighth, tenth, and twelfth powers we have the average values

$$\frac{151}{315}, \quad \frac{15619}{36288} \text{ and } \frac{655177}{1663200}. \qquad \ldots\ldots(4{\cdot}3)$$

The general term for this average can be written

$$\frac{1}{2}\cdot\frac{1}{(2r-1)!}\,\delta^{2r}u_0,$$

where $u_n=|n|^{2r-1}$.

Alternatively the actual values for particular points of interpolation, and not merely the averages over all points, of the mean powers of the errors may be reached by noting that the series of coefficients of the independent tabular errors, in the error of the interpolate, is

$$\frac{\sin\theta}{\theta}, \quad \frac{\sin(\theta\pm\pi)}{\theta\pm\pi}, \ldots\ldots,$$

so that the sum of their squares is 1 if $\theta=0$, and

$$\frac{1}{\pi^2}\sin^2\theta\,\mathop{S}_{r=-\infty}^{\infty}\left(\frac{1}{x+r}\right)^2, \quad \theta\neq 0,$$

where

$$x=\frac{\theta}{\pi},$$

Digitized by Google

and r is an integer. We have

$$\sum_{r=0}^{\infty} \frac{1}{(x+r)^2} = \frac{\partial^2}{\partial x^2} \log (x-1)!.$$

Also

$$\sum_{r=-\infty}^{-1} \frac{1}{(x+r)^2} = \frac{\partial^2}{\partial x^2} \log (-x)!.$$

[$(x-1)!$ is written for $\Gamma(x)$ even though x is not an integer.] Adding, we have

$$\sum_{-\infty}^{\infty} \frac{1}{(x+r)^2} = \frac{\partial^2}{\partial x^2} \log \frac{\pi}{\sin \pi x} = \pi^2 \operatorname{cosec}^2 \theta.$$

Thus the sum of squares is unity for all values of θ. The sum of the fourth powers is, similarly,

$$\frac{1}{3!} \sin^4 \theta \frac{\partial^2}{\partial \theta^2} (\operatorname{cosec}^2 \theta) = 1 - \tfrac{2}{3} \sin^2 \theta. \qquad \ldots\ldots(4{\cdot}4)$$

For the sixth powers we have

$$1 - \sin^2 \theta + \tfrac{2}{15} \sin^4 \theta, \qquad \ldots\ldots(4{\cdot}5)$$

and for the eighth powers

$$1 - \tfrac{4}{3} \sin^2 \theta + \tfrac{2}{5} \sin^4 \theta - \tfrac{4}{315} \sin^6 \theta. \qquad \ldots\ldots(4{\cdot}6)$$

These expressions are the coefficients of h^2, h^4, etc., in the expansion of

$$hx \cot (\theta - hx),$$

where $x = \sin \theta$.

The assumption is made that the error distribution of the interpolate will be symmetrical, so that only even moments need be considered. The two cases that arise are (*a*) normal distribution of tabular errors, which occur when the table is used as constructed, and not cut down, (*b*) rectangular distribution of errors, as in the case of all ordinary tables, which are correct to the last figure. So that in (*b*) if dp represent the small element of frequency ($= f(x)\,dx$), we have $dp = dx$ for all values of x between $-\frac{1}{2}$ and $+\frac{1}{2}$, say, and $dp = 0$ outside these limits.

The nature of the distributions is best exhibited by studying their characteristic functions. If, in the usual notation,

$$\mu_r' = \int_{-\infty}^{\infty} x^r dp$$

is the rth moment of the distribution (taken as unlimited in range) about an arbitrary origin, and μ_r is the corresponding

Digitized by Google

moment about the mean, then the Moment Generating Function is defined as

$$M = \int_{-\infty}^{\infty} e^{xt} dp = 1 + \mu_1' t + \mu_2' \frac{t^2}{2!} + \mu_3' \frac{t^3}{3!} + \ldots\ldots$$

$$= e^{\mu_1' t}\left(1 + \mu_2 \frac{t^2}{2!} + \mu_3 \frac{t^3}{3!} + \ldots\ldots\right). \quad \ldots\ldots(4{\cdot}7)$$

The logarithm of M may be called the Kappa Generating Function, and we have

$$\mathrm{K} = \log M = \kappa_1 t + \kappa_2 \frac{t^2}{2!} + \kappa_3 \frac{t^3}{3!} + \ldots\ldots, \quad \ldots\ldots(4{\cdot}8)$$

where the κ's are the semi-invariants of Thiele. Thus

$$\kappa_1 = \mu_1', \quad \kappa_2 = \mu_2, \quad \kappa_3 = \mu_3, \quad \kappa_4 = \mu_4 - 3\mu_2^2, \quad \kappa_5 = \mu_5 - 10\mu_2\mu_3, \text{ etc.}$$

An important property of the K-functions is that if the distribution of ϵ has for its characteristic function

$$\mathrm{K} = \kappa_1 t + \kappa_2 \frac{t^2}{2!} + \kappa_3 \frac{t^3}{3!} + \ldots\ldots,$$

then that of $p\epsilon$, where p is constant, has

$$\mathrm{K}_1 = p\kappa_1 t + p^2\kappa_2 \frac{t^2}{2!} + p^3\kappa_3 \frac{t^3}{3!} + \ldots\ldots$$

Also for any linear compound of independent variates the K-functions are added, i.e. for

$$X = p_1\epsilon_1 + p_2\epsilon_2 + p_3\epsilon_3 + \ldots\ldots,$$

where $p_1, p_2, p_3, \ldots$ are constant and $\epsilon_1, \epsilon_2, \epsilon_3, \ldots$ are independent variates of a distribution given by K, then for the distribution X

$$\mathrm{K}_2 = \kappa_1 S(p)\, t + \kappa_2 S(p^2) \frac{t^2}{2!} + \kappa_3 S(p^3) \frac{t^3}{3!} + \ldots\ldots \quad \ldots(4{\cdot}9)$$

This property enables us to reach very quickly the form of the distribution of the error of the interpolate for a given distribution of tabular errors. Thus for the rectangular distribution of common occurrence

$$M' = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{t'x} dx = \frac{2}{t'} \sinh \frac{t'}{2}$$

$$= 1 + \frac{1}{3 \cdot 2^2} \cdot \frac{t'^2}{2!} + \frac{1}{5 \cdot 2^4} \cdot \frac{t'^4}{4!} + \frac{1}{7 \cdot 2^6} \cdot \frac{t'^6}{6!} + \ldots\ldots$$

Adjusting so as to obtain a unit standard error by writing

$$t' = t\sqrt{12},$$

we have

$$M = 1 + \frac{t^2}{2!} + \frac{3^2}{5} \cdot \frac{t^4}{4!} + \frac{3^3}{7} \cdot \frac{t^6}{6!} + \ldots\ldots \quad \ldots(4{\cdot}10)$$

Digitized by Google

Then $$K = \frac{t^2}{2!} - \frac{6}{5}\cdot\frac{t^4}{4!} + \frac{48}{7}\cdot\frac{t^6}{6!} - \frac{432}{5}\cdot\frac{t^8}{8!} + \ldots\ldots \quad \ldots(4{\cdot}11).$$

So for the interpolate, using the results (4·4), (4·5), (4·6), we have

$$K = \frac{t^2}{2!} - \frac{6}{5}\left(1 - \frac{2}{3}\sin^2\theta\right)\frac{t^4}{4!} + \frac{48}{7}\left(1 - \sin^2\theta + \frac{2}{15}\sin^4\theta\right)\frac{t^6}{6!}$$
$$- \frac{432}{5}\left(1 - \frac{4}{3}\sin^2\theta + \frac{2}{5}\sin^4\theta - \frac{4}{315}\sin^6\theta\right)\frac{t^8}{8!} + \ldots. \quad \ldots(4{\cdot}12)$$

If the successive terms are averaged, by the use of (4·1), (4·2), (4·3), we have what may be called the synthetic distribution of the interpolate, given by the function

$$K = \frac{t^2}{2!} - \frac{4}{5}\cdot\frac{t^4}{4!} + \frac{132}{35}\cdot\frac{t^6}{6!} - \frac{7248}{175}\cdot\frac{t^8}{8!} + \ldots, \quad \ldots(4{\cdot}13)$$

in which the terms beyond the first indicate the departure from normality. It will be seen that the distribution does not tend to normality, even though the number of tabular errors involved is increased without limit. Pearson's β_2 tends to 2·2 instead of 3·0. If the moments are wanted the procedure is reversed and we obtain

$$M = 1 + \frac{t^2}{2!} + \frac{11}{5}\cdot\frac{t^4}{4!} + \frac{237}{35}\cdot\frac{t^6}{6!} + \frac{4127}{175}\cdot\frac{t^8}{8!} + \ldots. \quad \ldots(4{\cdot}14)$$

Alternatively we may consider the distribution compounded of the several distributions of the successive tabular errors. Thus if the frequency distributions are

$$dp = \phi_1 dx, \quad \phi_2 dx, \quad \phi_3 dx,$$

and so on, we consider the distribution $dp = \bar{\phi}dx$, $\bar{\phi}$ being the mean of $\phi_1, \phi_2, \phi_3, \ldots\ldots$ The corresponding Moment Generating Function is

$$M = \int_{-\infty}^{\infty} e^{xt}\bar{\phi}\,dx = \bar{M}.$$

This procedure is equivalent to reversing equation (4·12) and then averaging the terms. Thus

$$M = 1 + \frac{t^2}{2!} + \frac{9 + 4\sin^2\theta}{5}\cdot\frac{t^4}{4!} + \frac{135 + 180\sin^2\theta + 32\sin^4\theta}{35}\cdot\frac{t^6}{6!}$$
$$+ \frac{1575 + 4200\sin^2\theta + 2352\sin^4\theta + 192\sin^6\theta}{175}\cdot\frac{t^8}{8!} + \ldots\ldots$$
$$\ldots\ldots(4{\cdot}15)$$
$$= 1 + \frac{t^2}{2!} + \frac{11}{5}\cdot\frac{t^4}{4!} + \frac{237}{35}\cdot\frac{t^6}{6!} + \frac{4617}{175}\cdot\frac{t^8}{8!} + \ldots\ldots \quad \ldots(4{\cdot}16)$$

Digitized by Google

The synthetic and compounded distributions agree, as was to be expected from the constancy of the variance, up to the sixth moment.

5. The point we now make is that if the distribution of tabular errors is normal, as may be reasonably assumed for a table not cut down, the only term left in the K-function is $\frac{t^2}{2!}$. It follows that K for the error in the interpolate is always $\frac{t^2}{2!}$ for all values of θ, hence the distribution of error at any point of interpolation is exactly the same as that of the tabular error. The distinction between the synthetic and the compounded distributions now disappears.

A distribution of error such as that given by (4·12) or (4·15) is by no means easy to deal with. Variation will take place as the interval between the tabular entries is traversed, and even when this is averaged over the interval, and very high order interpolation formulae are used, the distribution is still of a type which has not been investigated, and for which probability integral tables are not available. The best that could be done would perhaps be to use the Pearsonian Type II

$$df = \frac{\frac{n-2}{2}!}{\sqrt{n\pi}\,\frac{n-3}{2}!}\left(1-\frac{x^2}{n}\right)^{\frac{n-3}{2}}dx, \quad n=\frac{11}{2},$$

which agrees with it as far as the fourth moment, and does not differ widely in the neighbouring higher moments. It is, however, a distribution of finite range, whereas the true range must tend logarithmically to infinity as the order of interpolation is increased.

On the other hand, when the distribution of tabular errors is normal, that of the interpolate is also normal, and is constant throughout the interval. The restriction still applies, however, that this result has only been demonstrated for high order interpolation. With formulae using fourth or sixth differences the distribution is still normal but with a slightly inconstant variance; the synthetic distribution formed by averaging the variances is normal, and may be used for most purposes. The convenience of doing so should be borne in mind when new tables are in contemplation, and it is suggested in particular that, before a manuscript table is cut down to ensure the accuracy of the last figure, the effect of this procedure on the accuracy of the interpolate should receive very careful consideration.

Digitized by Google

REFERENCES.

(1) A. J. THOMPSON (1921). *Table of the Coefficients of Everett's Central Difference Interpolation Formula. Tracts for Computers*, V. Cambridge University Press.

(2) —— (1924). *Logarithmetica Britannica*, Part IX. Issued by the Biometric Laboratory.

(3) KARL PEARSON (1920). *On the Construction of Tables and on Interpolation. Tracts for Computers*, II. Part I. Cambridge University Press.

(4) J. F. STEFFENSEN (1927). *Interpolation*. Baltimore: The Williams and Wilkins Company.

Digitized by Google

A differential retarding potential method for the study of the energy distribution of slow electron emissions. By C. F. SHARMAN, B.A., King's College. (Communicated by Professor Sir ERNEST RUTHERFORD.)

[*Received* 15 August, *read* 25 July 1927.]

In a previous paper* the relative merits of the two general methods for the determination of electronic energy distributions have been discussed in general. Both methods present special difficulties when a detailed study of the low energy region is attempted. In the first place, residual gas in the apparatus may, for obvious reasons, seriously distort the magnetic spectrum; whereas retarding potential measurements are much less affected since elastic collisions with molecules of the gas (which are a large fraction of the total number of collisions at these low energies) do not necessarily alter the ultimate destination of an electron. Secondly, an accurate knowledge of the intensity of the magnetic field becomes very necessary in this region, owing to the nature of the transformation required to obtain the energy distribution from the experimental results. On the other hand, the great advantage of the magnetic method remains, namely, its differential property.

The arrangement to be described, since it makes use of retarding potentials, does not suffer from the first two defects yet gives its results in the convenient differential form and needs no drastic transformation. It actually differentiates automatically the ordinary retarding potential curves. The method was applied to the secondary electron emission from a polished copper surface, excited by an electron beam of a few hundred volts energy.

Suppose that $$dn = A f(v)\, dv$$

represents the energy distribution to be studied, where A is a constant depending on the intensity of the primary beam and the nature of the target, f is a function determining the shape of the distribution curve, and dn is the number of electrons with energies in the range v to $v + dv$. The ordinary retarding potential method measures the current

$$\int_v^\infty A f(v)\, dv$$

represented by the area ABC of Fig. 1 (a).

* *Camb. Phil. Soc. Proc.* XXIII, Pt. 5 (1927), p. 523.

Digitized by Google

The original idea was to have two retarding potential arrangements, exactly similar in every respect, to apply a retarding potential v to one and $v+dv$ to the other, and to measure the *difference* of the currents to the two collecting electrodes. This would be represented in Fig. 1 (*a*) by the element of area *ABED*.

The conditions necessary for the application of this method are that both the constant A and the function f shall be identical for the two arrangements; that is, that the two distribution curves shall be exactly superposed. We can assume that for two targets carefully prepared under the same conditions the f-condition will be satisfied. The equality of the A's can only be effected in the present case by having two filaments to supply the primary beams, heated by the same battery, but separately adjustable. This was tried, but it was found that slight variations of one filament against the other entirely masked the small differential current, unless

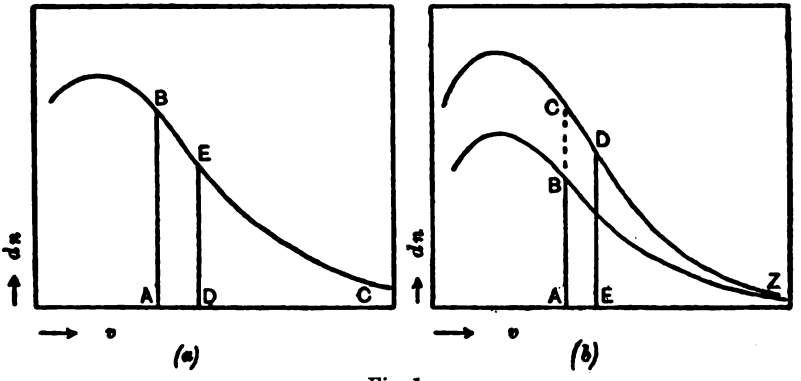


Fig. 1.

very large elementary potential differences dv were used. The "resolution" obtained was far too poor. In order to remove this difficulty, both primary beams were derived from a single filament and the balancing of the galvanometer was done artificially.

Fig. 2 shows the part of the apparatus in the tube, and is symmetrical about a plane through the filament and perpendicular to the axis of the tube.

Two beams of thermions were accelerated through the coarse grids *AA* of parallel vertical wires, and were defined by the 2 mm. circular holes *BB*. The filament was raised to a negative potential V and the steel cylinder containing the grids and holes was earthed through a milliammeter. The primary beams then passed through 3 mm. holes in the hemispherical collectors *CC* and struck the copper targets *TT* which were supported on glass stalks from the back plate of the collectors. The inner surfaces

Digitized by Google

of the collectors were smoked with a tallow candle and then heated in a Bunsen flame to free them from condensed hydrocarbons. In these experiments the dimensions of the collectors (35 mm. in

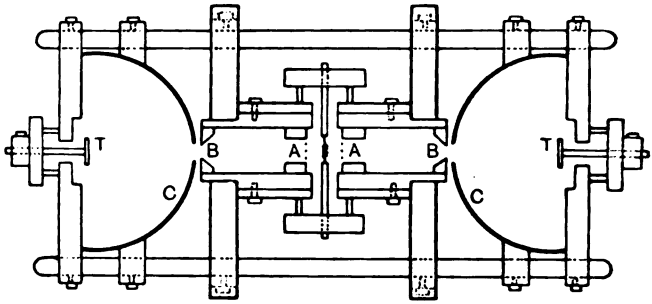


Fig. 2.

diameter) were limited by the size of the enveloping glass tube; their efficiency would have been improved by using complete spheres of very much greater radius. The collectors were supported from the two large central discs by three glass rods set into the various discs with screws. The alignment of the whole was easily verified by removing the targets and looking along the axis.

The electrical connections are shown diagrammatically in Fig. 3.

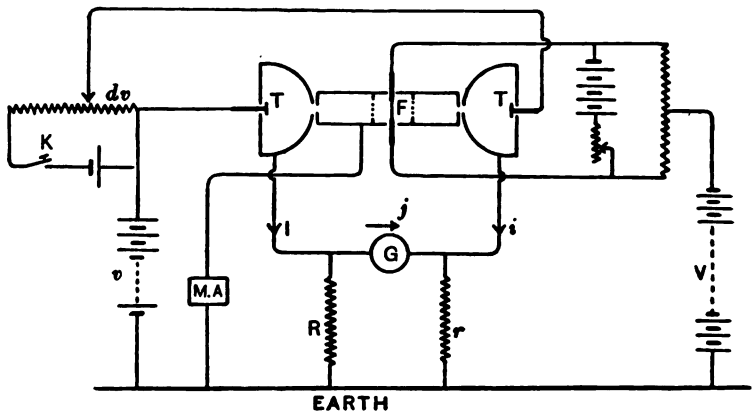


Fig. 3.

The secondary electrons were retarded by charging the targets positively, the collectors remaining very nearly at zero potential. This, of course, increases the energy of the primary beams on reaching the targets, but owing to the small range of the retarding potentials and to its greater convenience, this method was used in

Digitized by Google

preference to charging the collectors. Both targets were charged by the same battery v, but on one side was included an insulated potentiometer, which on the completion of its circuit by the key K, added a small extra retarding potential dv. Generally this was 0·1 volt.

The arrangement for measuring the difference of the two currents is shown in the diagram. The currents passed to earth through two resistances r and R, and the difference of potential between their ends was measured by a sensitive galvanometer (sensitivity 10^{-10} amp.). It is necessary that the difference of potential across the galvanometer shall be small compared with dv, and that those across r and R shall be small compared with v. These conditions were easily satisfied in practice.

Now suppose that, for the same retarding potential on the two sides, the currents are given by

$$\left.\begin{aligned} i &= a\int_v^\infty f(v)\,dv \\ I &= A\int_v^\infty f(v)\,dv \end{aligned}\right\} \quad \text{.....(1),}$$

being represented by areas BAZ and CAZ respectively of Fig. 1 (b).

The resistances r and R are now adjusted to give no current in the galvanometer. Hence, for $dv=0$, we have $j=0$, and

$$ir = IR \quad \text{.....(2).}$$

On applying the extra retarding potential dv, suppose that the current I is changed to $I+dI$. Then, by Kirchoff's Laws,

$$(I+dI-j)\,R-(i+j)\,r = j\,G \quad \text{.....(3)}$$

and, substituting from (2),

$$dI = j\,(G+R+r)/R \quad \text{.....(4).}$$

But $$dI = Af(v)\,dv.$$

Hence, for this particular value of v, the galvanometer current j and the distribution function $f(v)$ are connected by the relation

$$f(v)\,dv = j\,(G+R+r)/RA.$$

Now from equations (1) and (2) we see that

$$i/I = a/A = R/r \quad \text{.....(5).}$$

Since this fraction is independent of v, the adjustment of r and R, once made, holds good for the whole range of v; and j plotted against v gives the form of the energy distribution function.

The procedure, then, was as follows: the resistances r and R were adjusted until for $v=0$ and $dv=0$, the galvanometer was

Digitized by Google

undeflected*. The reading was taken on closing the key K. This was repeated for various values of v in the range -10 to $+20$ volts. If relation (5) is correct, the zero reading should remain unchanged as v is varied. Actually there was a small movement, due probably to an inevitable slight difference in the targets; but this was never serious enough to require a readjustment of r and R in any one run. The targets were cut from the same sample of copper, and were carefully polished on the same polishing agent. The total filament emission was read on the milliammeter before every galvanometer reading.

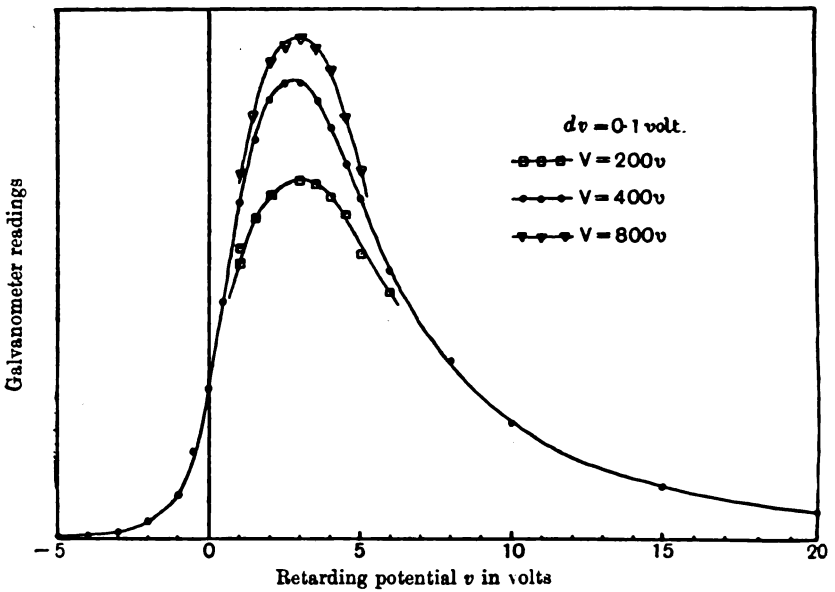


Fig. 4.

A typical curve is shown (black dots) in Fig. 4 for a primary energy of $400v$. The portions of curves indicated by squares and triangles were obtained with accelerating voltages of $200v$ and $800v$ respectively. Except at the peaks they are very nearly coincident. Two facts emerge from the study of these curves. In the first place, the position of the peaks and the general shape do not appear to depend on the primary energy: this has been suggested before by Becker† for a primary electron beam, and by

* The actual magnitudes of r and R are governed in practice by the relations

$$ir = IR \quad \text{and} \quad \bar{R} = (G + R + r)$$

where $\bar{R}$ is the resistance for critical damping of the galvanometer.

† Becker, *Ann. der Physik*, LXXVIII (1925), p. 209.

Digitized by Google

Becker* and Campbell† for a primary beam of alpha particles, both from consideration of the ordinary retarding potential results. Secondly, so far as the low energy region is concerned, the intensity of secondary emission increases with the primary energy, although much more slowly between $400v$ and $800v$ than between $200v$ and $400v$. This conclusion is not necessarily incompatible with the results of many workers‡ who have found a maximum of the "coefficient of secondary emission" at about $300v$, since they have been dealing with the total unresolved emission, corresponding to the area under these curves.

It will be noticed that the current through the galvanometer did not fall to zero in passing through the origin, which would appear to indicate electrons of negative energy. This effect, which was expected, is of course instrumental, and its explanation is easily found. Consider the conditions inside a collector with a retarding potential v in operation. The collector will receive all electrons from the target with emission energies greater than v, but of these a small fraction (depending on the energy with which they impinge on the collector) will be "reflected." And of these again some will be received back by the target and so will be lost from the point of view of the collector. Hence each collector will not receive its total current (which it should get when $v = 0$) until the applied field is reversed and is intense enough to prevent reflected secondaries from returning to the target. This latter fraction depends in a positive manner on the solid angle subtended by the target at a point on the collector. In the present arrangement, as previously explained, this solid angle was rather large; but even so, the effect was not great enough to mask the general characteristics of the energy distribution in this region. There is no doubt of the existence of the maximum—in this case at about 3 volts; and since all ordinates on the low energy side must be reduced slightly to correct for reflection from the collectors, there is a strong indication that the true distribution curve passes through the origin.

It was noticed that the position of the maximum moved as the targets were bombarded. A series of distribution curves was therefore taken with a constant primary energy ($200v$) and a constant total filament emission, beginning with clean highly-polished targets. The whole run lasted fifteen hours and was done on three successive days. The circles in Fig. 5 show how the position of the peak varied with the time the targets had been

* Becker, *Ann. der Physik*, LXXV (1924), p. 217.

† Campbell, *Phil. Mag.* XXIII (1914), p. 46.

‡ Petry, *Phys. Rev.* 1925, p. 346; Farnsworth, *Phys. Rev.* XXVI (1925), p. 41; Barber, *Phys. Rev.* 1921, p. 322.

Digitized by Google

bombarded, the scale of position of the peak being in volts on the right.

In the same figure is plotted the corresponding height of the peak in galvanometer deflections. The first pair of points is less reliable than the rest, since it corresponds to a curve taken hurriedly immediately at the start, with $dv = 0{\cdot}04v$, whereas for all the other points $dv = 0{\cdot}1v$.

An examination of the targets at the end of the run revealed circular patches in the middle, of a size corresponding to the geometrical image of the filament and slit system, and of a deep purple colour. They were undoubtedly films of tungsten formed by evaporated tungsten molecules which had travelled in free paths from the filament to the targets*. The deposition of these films was followed in a separate experiment and the sequence of

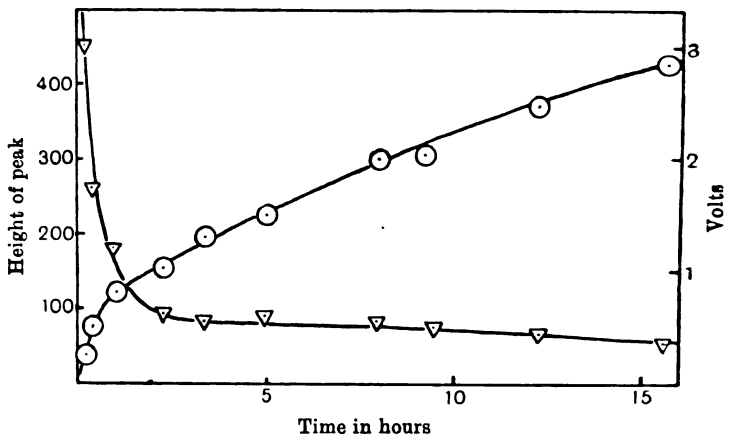


Fig. 5.

colour changes was as follows: first a gradually deepening brown coloration which turned to a bright red, then followed a number of alternations from bright red to bright green, then a gradually deepening purple and finally black. Mr J. D. Cockcroft, from a comparison with films of cadmium deposited by him at a known rate, has kindly given me his estimate of the thickness of the purple film—rather less than 20 atoms. This would correspond to the point on the extreme right of Fig. 5. On this rough estimate, and assuming that the deposit was laid down at a uniform rate, it is seen from both curves of Fig. 5 that there is an indication of a change of secondary emission characteristics when the film reaches

* The sharpness of these images showed that the vacuum conditions inside the experimental tube were adequate.

Digitized by Google

the monatomic stage (that is after one hour in this experiment). After this the distribution changes much more slowly, but still appreciably.

These curves seem to receive a reasonable explanation if the mechanism of secondary emission is considered as made up of two separate processes: (*a*) the excitation of the (free) electrons of the metal to energies large compared with the thermal energy corresponding to the temperature of the target, and (*b*) the escape of these through the surface. The rapid changes accompanying the deposition of the first monatomic layer would be due to the modification of the second process as the *surface only* changes from copper to tungsten. The slower and more sustained variation of the distribution must then be caused by the change from copper to tungsten of the deeper layers where the secondaries are actually excited. A very drastic extrapolation indicates that the position of the peak becomes constant at about 50 atoms, which may be taken as giving a rough idea as to the number of monatomic layers contributing to the secondary emission. If the above explanation is correct, these experiments give some useful information which would be difficult to obtain by any other means.

It seems likely that this contamination of targets by evaporated molecules from the filament supplying the primary beam must have entered into much of the work on secondary emission. It may account for some of the contradictory results which appear in the literature on the subject.

Summary.

The method described gives the differential energy distribution curve directly from readings of a galvanometer and a voltmeter. When adjusted, the arrangement is convenient in operation and very good "resolution" can be obtained. The existence of a maximum frequency of secondary electrons at a finite small energy has been established, and its position is shown to be independent of the energy of the primary beam over a fairly wide range. The effect of deposition of tungsten on a clean copper surface is studied with respect to the secondary emission properties of the surface. From these experiments some interesting conclusions concerning the mechanism of the process may be drawn.

I am very grateful to Prof. Sir Ernest Rutherford for the encouragement he has given and the interest he has taken in this work.

Digitized by Google

The emission of radiation by a quadripole electric moment on the Quantum Mechanics. By J. A. GAUNT, B.A., Trinity College, and W. H. MCCREA, B.A., Trinity College.

[*Received* 15 August, *read* 24 October 1927.]

The usual method of calculating the probability of a switch between stationary states by the Quantum Mechanics is really equivalent to finding the intensity of the dipole radiation. Other switches of the system not involved in this radiation may be possible but are not taken account of by the calculation. A method of calculating the probability of switches due to radiation by the quadripole moment seems to be supplied by Dirac's recent theory of the interaction between matter and radiation in his paper "The Quantum Theory of Dispersion*." Equations from this paper will be denoted by a D. The results are probably of a purely theoretical interest and the intensities found probably too faint to be observed; but the present considerations do appear to emphasise the need for including the radiation in the exact theoretical treatment of a Quantum Mechanical system.

The radiation field is given by its vector potential $(\kappa_x \kappa_y \kappa_z)$, chosen so that the scalar potential is zero. The classical relativistic Hamiltonian for an electron moving in this radiation and an electrostatic field of potential ϕ is:

$$\begin{aligned} H &= c\{[m^2c^2 + p_x^2 + p_y^2 + p_x^2] + 2e/c\,.\,(p_x\kappa_x + p_y\kappa_y + p_z\kappa_z) \\ &\qquad + e^2/c^2\,.\,(\kappa_x^2 + \kappa_y^2 + \kappa_z^2)\}^{\frac{1}{2}} - e\phi \\ &= c[m^2c^2 + p_x^2 + p_y^2 + p_z^2]^{\frac{1}{2}} - e\phi + e/c\,.\,(\dot{x}\kappa_x + \dot{y}\kappa_y + \dot{z}\kappa_z) \\ &\qquad + e^2/2mc^2\,.\,(\kappa_x^2 + \kappa_y^2 + \kappa_z^2) \\ &= H_0 + \Sigma e/c\,.\,\dot{x}\kappa_x + \Sigma e^2/2mc^2\,.\,\kappa_x^2 \qquad \text{.....(D 10)} \end{aligned}$$

where H_0 is the Hamiltonian in the absence of radiation. The largest terms neglected are proportional to $e\dot{x}^3\kappa_x/c^3$, $e^2\dot{x}^3\kappa_x^2/mc^4$.

The second term of the Hamiltonian (D 10) will express the interaction between the radiation and any charged particle, if e is suitably interpreted; it will also apply to several particles if the summation is taken over the coordinates of each of them.

As a rule, κ_x is given its value at the centre of the system. The interaction term is then proportional to $\Sigma e\dot{x}$, the rate of change of the dipole moment. This gives rise to the usual formula for the intensity (see Dirac). When, however, the dipole moment vanishes,

* *Proc. Roy. Soc.* A, 114 (1927), pp. 710–728.

Digitized by Google

it is necessary to go a step further, and to add to the central value of κ_x the next term in its Taylor's expansion, viz.:

$$\left(x\frac{\partial}{\partial x}+y\frac{\partial}{\partial y}+z\frac{\partial}{\partial z}\right)\kappa_x \quad \text{.....(1).}$$

The derivatives in (1) are to be evaluated at the origin.

Now if the radiation is analysed into plane-polarised monochromatic plane waves, the contribution to the vector potential of the rth component is

$$\kappa_r = 2\,(h\nu_r/2\pi c\,\sigma_r)^{\frac{1}{2}}\,N_r^{\frac{1}{2}}\cos 2\pi\,[\theta_r/h-\nu_r\xi_r/c]$$

...(D 11 and p. 719).

Here h is Planck's constant, ν_r the frequency, θ_r/h the phase, N_r the number of quanta, σ_r the number of components per unit solid angle per unit frequency, and ξ_r is the component in the direction of propagation of the vector (xyz). Thus, (1) is equal to

$$\sum_r \xi_r\,.\,2\,(2\pi h\,\nu_r^3/c^3\sigma_r)^{\frac{1}{2}}\,N_r^{\frac{1}{2}}\sin 2\pi\theta_r/h\,.\cos\alpha_{xr} \quad \text{.....(2)}$$

where α_{xr} is the angle between the x-axis and the direction of κ_r, which is the direction of the electric vector of the rth component. Substituting (2) for κ_x in the second term of (D 10), and using η_r for the projection of (xyz) on the electric vector of the rth component, we have for the interaction term in the Hamiltonian

$$\sum_r\Sigma e/c\,.\,\xi_r\dot{\eta}_r\,.\,2\,(2\pi h\,\nu_r^3/c^3\sigma_r)^{\frac{1}{2}}\,N_r^{\frac{1}{2}}\sin 2\pi\theta_r/h \quad \text{.....(3).}$$

The second summation is over the various particles.

It should be pointed out that (3) is larger than the terms neglected in (D 10), at any rate for frequencies such as are emitted by the system. For it is proportional to $e\dot{x}x\nu\kappa_x/c^2 \propto e\dot{x}^2\kappa_x/c^2$.

So far the calculation has been classical, and the order of factors has been irrelevant. In quantum theory, however, N_r and θ_r are non-commuting conjugate variables, while ξ_r, $\dot{\eta}_r$ depend on the dynamical variables of the atom or molecule, and may not commute with each other, though they do with N_r and θ_r. It is necessary to introduce the "real q-numbers":

$$\frac{1}{2i}\{N_r^{\frac{1}{2}}\,e^{2\pi i\theta_r/h}-(N_r+1)^{\frac{1}{2}}\,e^{-2\pi i\theta_r/h}\} \quad \text{for } N_r^{\frac{1}{2}}\sin 2\pi\,\theta_r/h,$$

$$\tfrac{1}{2}(\xi_r\dot{\eta}_r+\dot{\eta}_r\xi_r) \quad \text{for } \xi_r\dot{\eta}_r.$$

The interaction term (3) becomes

$$\sum_r\Sigma e/2ic\,.\,(2\pi h\,\nu_r^3/c^3\sigma_r)^{\frac{1}{2}}\,(\xi_r\dot{\eta}_r+\dot{\eta}_r\xi_r)$$
$$\{N_r^{\frac{1}{2}}\,e^{2\pi i\theta_r/h}-(N_r+1)^{\frac{1}{2}}\,e^{-2\pi i\theta_r/h}\} \quad \text{...(4).}$$

Digitized by Google

This may be compared with the third term in (D 13), which expresses the dipole interaction. The deduction of the intensity of the radiation is exactly parallel to Dirac's argument, and the result is an obvious adaptation of Dirac's on p. 719.

For a switch from a state denoted by J' to one denoted by J'', with the emission of the appropriate quantum of radiation, the value of Einstein's A coefficient per unit solid angle for each plane of polarisation is

$$2\pi^3 e^2/hc^5 \cdot \nu^3 (J'J'') \, |\Sigma (\xi_r \dot{\eta}_r + \dot{\eta}_r \xi_r)(J'J'')|^2 \qquad \ldots(5).$$

We can obtain in this way the quadripole radiation of a Planck oscillator. We suppose that the charge e is vibrating about the origin along the x-axis and have to evaluate $(\xi\dot{\eta} + \dot{\eta}\xi)(J'J'')$ [dropping the r's now since we are dealing with the emission of a single quantum]. Let the direction of propagation make an angle α with Ox and take the plane of polarisation through this direction and Ox. The component in the perpendicular plane will be zero since η is then zero.

We have

$$(\xi\dot{\eta} + \dot{\eta}\xi)(J'J'') = 2\pi i \sum_{J'''} [x(J'J''')\, x(J'''J'')\, \nu(J'''J'') + x(J'J''')\, \nu(J'J''')\, x(J'''J'')] \sin\alpha \cos\alpha \qquad \ldots(6).$$

Now for a simple harmonic oscillator*

$$x(n, n') = 0 \text{ unless } n' = n \pm 1,$$

$$x(n, n-1) = \frac{1}{2\pi}\sqrt{\frac{nh}{2\nu_0 m}},$$

$$\nu(n, n-1) = \nu_0,$$

where ν_0 is the characteristic frequency and m the mass of the particle.

Thus the quantity in (6) is not zero only if the switch is $n \to n$ (no radiation) or $n \to n \pm 2$. The A coefficient for $n \to n-2$ is, from (5), there being no contribution from the dipole radiation,

$$\pi \frac{e^2 h \nu_0^3}{m^2 c^5} n(n-1) \sin^2 2\alpha,$$

and the frequency is, of course, $2\nu_0$.

The factor $\sin^2 2\alpha$ gives the distribution of the radiation in space; it vanishes along the axis and in the plane through O perpendicular to it.

Thus there is a finite probability of a switch $\Delta n = 2$. There will be the higher switches as well but our original Hamiltonian (D 10) would have to be replaced by a higher approximation to obtain them.

* Schrödinger, *Ann. der Phys.* 79, p. 514 (1926).

Digitized by Google

To find the radiation from a rigid homopolar molecule, possessing equal charges e separated by a distance $2l$, rotating about its centre, suppose its position defined by Eulerian angles θ, ϕ, ψ so chosen that the electric vector is in the direction $\theta = 0$ and the direction of propagation lies in the plane $\phi = 0$. Then

$$\xi = l \sin\theta \cos\phi, \qquad \eta = l \cos\theta$$

and we have to calculate

$$\sum_{J'''} [(\sin\theta\cos\phi)(J'J''')\cos\theta(J'''J'')\nu(J'''J'') + \cos\theta(J'J''')\nu(J'J''')(\sin\theta\cos\phi)(J'''J'')].$$

Summing for the two particles just introduces a factor 2.

The matrix components involved are amongst those calculated by Rademacher and Reiche*. It might be remarked, however, that their working can be shortened by using a form of the eigenfunction for a symmetrical rotator given by Darwin†. We shall merely write down the values for this expression when it is non-zero and when $J'' \neq J'$. We denote it by $\{r', s', t'; r'', s'', t''\}$ where r, s, t are the three quantum numbers. The possibility of axial spin is included; it exists when $t \neq 0$.

$$\{r, s, t;\ r-1, s-1, t\} = \frac{h}{4\pi^2 A} t \left(\frac{r-s+1}{r(r+1)} + \frac{r-s}{(r-1)r}\right) \sqrt{\frac{(r+s)(r+s-1)(r-t)(r+t)}{(2r+1)(2r-1)}},$$

$$\{r, s, t; r-2, s-1, t\} = \frac{h}{4\pi^2 A} \left(\frac{1}{r} + \frac{1}{r-1}\right) \frac{1}{2r-1} \sqrt{\frac{(r+s)(r+s-1)(r+s-2)(r-s)(r+t)(r+t-1)(r-t)(r-t-1)}{(2r+1)(2r-3)}},$$

where A is the moment of inertia about an axis perpendicular to the figure axis.

There will be similar, though not identical, results for

$$\{r, s, t; r-1, s+1, t\} \text{ and } \{r, s, t; r-2, s+1, t\}.$$

If we call the eigenfunction corresponding to a state (r, s, t) $\psi(r, s, t)$, then $\psi(r, s, t)$ and $\psi(r, s, -t)$ belong to states of the same energy. Following Hund‡ we can use instead the pair of eigenfunctions $\frac{1}{\sqrt{2}}\psi(r, s, t) \pm \frac{1}{\sqrt{2}}\psi(r, s, -t)$. Each of these, from the properties of hypergeometric polynomials, is alternately symmetrical and antisymmetrical as r takes successive integral values and one is symmetrical when the other is antisymmetrical. The

* *Zs. f. Phys.* 41, 1927, p. 453.
† *Proc. Roy. Soc.* A, 115, p. 1 (1927); see p. 5.
‡ *Zs. f. Phys.* 43, p. 805 (1927) (see p. 821).

Digitized by Google

corresponding states can only combine accordingly, symmetrical with symmetrical, antisymmetrical with antisymmetrical. We see that our results are in exact agreement with this since

$$\{r, s, t;\ r-1, s-1, t\}$$

changes sign with t while $\{r, s, t;\ r-2, s-1, t\}$ is unchanged when the sign of t is changed*.

It will be noticed that the switch $r \rightarrow r-1$ has non-zero intensity only if $t \neq 0$. This corresponds to the fact that the spin about the figure axis gives directional properties to the molecule, so that it does not, so to speak, repeat itself at every half revolution about the perpendicular axis.

The quantum number t can never change in a quantum switch, which is what we should expect. We can, however, apply similar working to the case of a symmetrical rotator with an electric moment perpendicular to its axis and obtain selection rules for t.

* For example $[\psi(r, s, t) + \psi(r, s, -t)] \rightarrow [\psi(r-1, s-1, t) - \psi(r-1, s-1, -t)]$ gives $2\{r, s, t;\ r-1, s-1, t\}$, the cross terms $\psi(r, s, t) \rightarrow \psi(r-1, s-1, -t)$, etc. giving zero. The factor 2 disappears when we normalise.

Digitized by Google

The Amount of Energy emitted in the γ-Ray Form by Radium E. By G. H. Aston, B.A., Trinity Hall. (Communicated by Prof. Sir E. Rutherford.)

[*Received* 19 August, *read* 24 October 1927.]

The β-ray type of disintegration is accompanied in general by two forms of radiation, the β-rays and the γ-rays. Radium E is noteworthy in that it gives a large amount of β-radiation but so little γ-radiation that it can only be detected by careful measurements. Yet, since it differs in this way from β-ray elements such as radium B and radium C, it is important for theoretical discussion that we should obtain as much knowledge as possible about this small amount of γ-radiation.

The first characteristic feature of γ-rays to be determined is their quality, i.e. a rough knowledge of the wave-lengths, and this has been done several times in the case of radium E by the usual method of finding the absorption coefficient. The results show that these γ-rays lie between the X-ray region and the hard γ-rays of radium C. For many radioactive elements precise determinations of the wave-lengths have been made by measuring the energy of the β-ray lines in the magnetic spectrum, these groups of homogeneous β-rays being produced by the ejection of electrons from the extranuclear part of the atom by the γ-rays from the nucleus. This method cannot be applied to radium E as no lines have been observed in the spectrum, either because they are too faint as a result of the weakness of the γ-rays, or possibly because, as is suggested later, the γ-rays may have a continuous spectrum and hence would not produce homogeneous β-rays.

The second important determination for γ-rays, and the one with which we are concerned chiefly in this paper, is the amount of energy emitted per disintegration. A knowledge of the effective wave-length gives an approximate mean value of the energy quantum, and thus the number of quanta per disintegration can be found. Though this energy determination is of considerable importance, it has only been carried out for radium B and radium C*. In this case it was found by measuring the heat produced on the absorption of the γ-rays, but considerable difficulty was experienced since the γ-ray energy is only a small part of the total emission. The γ-ray energy of radium E is much smaller, and also the available sources are much weaker than for radium B and radium C. The heating method is therefore impracticable for radium E. Fortunately, however, there remains a simple ionisation method, which, though not capable of giving a precise

* Ellis and Wooster, *Phil. Mag.* 50, p. 521, 1925.

Digitized by Google

value, yet determines limits within which the energy must lie. As there does not at present seem to be any method of measuring the energy accurately, it was thought that this method was of sufficient value to be carried out.

Principle of the Method.

There are two separate determinations involved in finding the γ-ray energy per disintegration. The total energy emitted in γ-ray form from a particular source has to be measured, and also the number of disintegrations occurring per second in this source. As pointed out above, a direct measurement of the energy is not practicable, and it was necessary to derive it from that of radium B and radium C by comparing ionisations. If the effective wave-length of radium E were the same as that of radium B or radium C, then the ratio of the ionisations would also be the ratio of the energy emissions. Since we have not this simple case, a knowledge of the variation of ionisation with wave-length is required to convert ionisation ratios into energy ratios. At present this has not been found, but the difficulty can be avoided by using simply the fact that the ionisation for γ-rays of given energy decreases as the wave-length decreases. The effective wave-length of radium E is longer than that of radium C and therefore by taking energy to be proportional to ionisation, we obtain a value of the energy of radium E in terms of that of radium C, which is too large. Similarly, considering radium B, for which the effective wave-length is longer than for radium E, we obtain a value for energy of radium E in terms of that of radium B which is too small. By using the ionisation ratios in this way we can obtain limits within which the radium E energy must lie.

It was possible to find the strength of the radium E in terms of the number of disintegrations per second by using a β-ray electroscope which had been calibrated by Dr Ellis and Mr Wooster during the course of their measurement of the heating effect of radium E*. The accuracy of the heating measurements was about 10°/$_{\circ}$ and the uncertainty in the strength of the radium E source used in the experiment described below is certainly not more than 20°/$_{\circ}$, and is probably less.

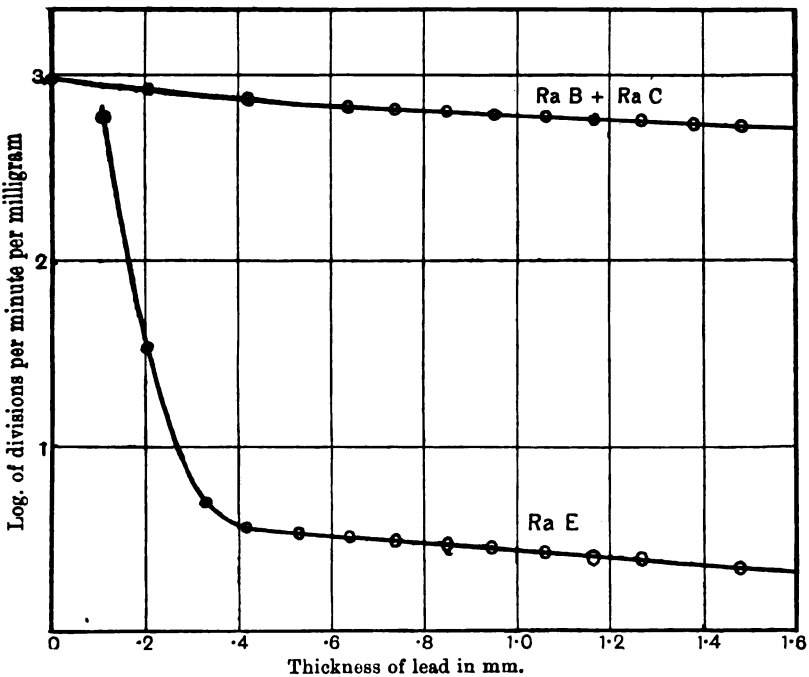
Experimental Details and Results.

With the amount of radium E that was available the ionisation by the γ-rays was small, and therefore a large electroscope of about 5 litres capacity was used. The sources were placed in a fixed position about 4 cms. below the base of the electroscope,

* *Nature*, April 16, 1927.

Digitized by Google

which consisted only of a thin aluminium foil having absorption negligible compared with that of the lead foils used. An absorption curve for radium E was found using thin lead foils, and then measurements were made under the same conditions, using a source of radium B and radium C in equilibrium. The source of radium E was obtained by electrolytic deposition on a platinum wire from a solution of radium D containing radium E and polonium in equilibrium. It was free from radium D, and the amount of polonium present was not more than a tenth of the radium E (in terms of the rate of disintegration), so that it would



not have any effect. The strength was found by using the calibrated β-ray electroscope mentioned above, and was ·54 mg. (i.e. it disintegrated at the same rate as this weight of radium) at the beginning of the experiment, and the measurements were corrected for the decay, using a half-period of 5·0 days. The radium B and radium C source consisted of radium bromide with its equilibrium products in a thin metal tube, its strength being ·22 mg. The thickness of this thin metal filter which alters the relative ionisation of radium B and radium C was not known, but it will be shown below that this does not invalidate the limits found for the radium E energy.

Digitized by Google

The graphs are drawn in each case for a milligram source. The kink in the radium E curve corresponds to the thickness of lead for which the β-particles are just stopped. This is 0·35 mm., giving a value of 0·40 grams per sq. cm. The corresponding value for aluminium was found by H. W. Schmidt* to be 0·47 grams per sq. cm. The β-particles from the radium B and radium C could not enter the electroscope as they were stopped by the metal cover of the source. The absorption coefficient of the γ-rays of radium E over a range ·5 to 1·5 mm. is 4·95 cm.$^{-1}$, i.e. $\frac{\mu}{\rho} = \cdot 43$. Schmidt found the absorption coefficient in aluminium over a similar range just past the complete absorption of the β-particles to be about ·9 cm.$^{-1}$, i.e. $\frac{\mu}{\rho} = \cdot 33$.

The graphs can be used to find the ratio of the ionisations of the two sources for any thickness of lead up to 1·5 mm. The results are calculated below for a thickness of 1·0 mm., since this is a convenient value in the middle of the graphs, and it also has bearing on the heating measurement of Ellis and Wooster mentioned above, as the walls of the calorimeter were of lead of this thickness and the calculation below gives the energy of the γ-rays not absorbed in the calorimeter.

For this thickness the ratio of the ionisation of radium E to that of radium B plus radium C for sources of equal strength is $4{\cdot}25 \times 10^{-3}$. We now require to divide up the ionisation between radium B and radium C. The measurements of Moseley and Makower† can be used; they find that after the radiation has passed through 1 mm. of lead about 16·5 % of the ionisation is due to radium B. Hence

$$\frac{\text{radium E ionisation}}{\text{radium B ionisation}} = 2{\cdot}57 \times 10^{-2}$$

and

$$\frac{\text{radium E ionisation}}{\text{radium C ionisation}} = 5{\cdot}08 \times 10^{-3}.$$

As mentioned above, the γ-rays of radium E are harder than those of radium B and softer than those of radium C, e.g. Curie and Fournier found $\frac{\mu}{\rho}$ in aluminium to be 0·092 after passing through 3 cms. of the absorber. The corresponding values of $\frac{\mu}{\rho}$ for radium B and radium C are 0·21 and 0·043. Therefore the radium E and radium B ionisation ratio gives a lower limit for

* *Phys. Zeit.* 8, p. 362, 1907.
† *Phil. Mag.* 23, p. 302, 1912.

Digitized by Google

the energy and the radium E and radium C ratio an upper limit. Ellis and Wooster* find the sum of the γ-ray energies of radium B and radium C to be $17{\cdot}0 \times 10^5$ volts per atom disintegrating†. Taking the ratio of the energies to be 0·085‡, the energies of radium B and radium C are respectively $1{\cdot}33 \times 10^5$ and $15{\cdot}7 \times 10^5$ volts per atom disintegrating. We require to deduce the energies after 1 mm. of lead. Most of the energy is contained in the high frequencies for which the absorption is small. Taking the absorption coefficient for radium C as ·5 cm.$^{-1}$, the energy passing through 1 mm. of lead is $14{\cdot}9 \times 10^5$ volts, and that for radium B as 3·0 cm.$^{-1}$, the energy $0{\cdot}98 \times 10^5$ volts. We then find that the γ-ray energy of radium E escaping through 1 mm. of lead lies between 2500 and 7600 volts per atom disintegrating. The effect of the thin metal covering of the radium B and radium C which has been neglected is to increase the ionisation of radium C relative to radium B. This correction would therefore bring the limits slightly closer together, so neglecting it does not affect the validity of the argument. The ratio of the energy of radium B to the energy of radium C is not known accurately, but an alteration of the value assumed for it would not alter appreciably the value of the energy of radium C though the value for the energy of radium B would be changed. Therefore the lower limit only is dependent on this division of energy.

It is of interest to obtain a value for the order of the energy by a different method. This consists in using the ratio of the β-ray and γ-ionisations, and making the assumption, which is true to within a factor of two or three, that for the same energy the ionisation of β- or γ-rays is proportional to the absorption coefficient. From Schmidt's§ measurements, the initial ratio of the ionisations is 6230, any soft γ-rays having an absorption coefficient comparable with that of the β-rays not being included in this value. The absorption coefficient in lead of the γ-rays is 4·95 cm.$^{-1}$, and therefore the ratio of the initial β-ray ionisation to the γ-ray ionisation after 1 mm. of lead is 10,200. The initial absorption coefficient of the β-rays is 315 cm.$^{-1}$. Making the assumption mentioned above, the ratio of the γ-ray energy after 1 mm. of lead to the β-ray energy is found to be $7{\cdot}1 \times 10^{-3}$. Ellis and Wooster find the mean β-ray energy to be 360,000 volts per atom disintegrating. The γ-ray energy after 1 mm. of lead is therefore about 2600 volts per atom disintegrating. The agreement with the value by the other method confirms the result that the γ-ray energy is very small and is of the order of 1% of the β-ray energy.

* *loc. cit.*
† Using the value $3{\cdot}7 \times 10^7$ atoms disintegrating per sec. per milligram.
‡ *Camb. Phil. Soc.* xxiii, p. 717, 1927. § *loc. cit.*

Digitized by Google

The most important use that can be made at present of the value for the energy per disintegration is to obtain a rough estimate of the number of quanta emitted per disintegration. We require for this purpose a value for the effective mean wave-length. This can be obtained from the X-ray result that the atomic photoelectric coefficient of absorption is about $2{\cdot}29 \times 10^{-2}\lambda^3 N^4$. The atomic coefficient of absorption of the γ-rays of radium E is $14{\cdot}8 \times 10^{-23}$. The part due to scattering must be subtracted and as this does not vary rapidly with the wave-length the value for radium B and radium C, $1{\cdot}76 \times 10^{-23}$, can be used. The photoelectric absorption coefficient is therefore $13{\cdot}0 \times 10^{-23}$ and the effective mean wave-length deduced from this is 0·050 Å.U. The quantum corresponding to this is about 250,000 volts. The true mean value for the quantum may, however, be higher than this if the γ-rays extend over a considerable range of wave-lengths, since the absorption coefficient is largely determined by the rays of longer wave-lengths. If we take for the γ-ray energy per disintegration after 1 mm. of lead the mean value of 6000 volts, then the initial energy will not be more than 10,000 volts, and probably will be less.

It follows that only one γ-ray quantum can be emitted during about thirty disintegrations of radium E atoms. This case is quite unlike that of radium B or radium C where roughly one quantum is emitted per disintegration. It seems probable that the origin of the γ-rays is different in the two cases. We know that most, if not all, of the γ-ray energy of radium B and radium C is in the form of a number of homogeneous rays, and there is some evidence for supposing that these are due to quantum changes between energy levels in the nucleus. Gray has shown that the β-rays of radium E falling on matter excite γ-rays* and probably the primary γ-rays are produced in a similar manner within the disintegrating atoms. The effective wave-length is considerably shorter than that of the K-radiation of radium E, and, taking into consideration the fact that no lines have been observed in the β-ray spectrum of radium E, it seems likely that these γ-rays have a continuous spectrum. Their origin is possibly to be found in the retardation or collision of the β-particles escaping from the nucleus. It has been suggested that the distribution of the velocities of the β-particles of a particular element over a wide range may be due to the β-rays originally homogeneous exciting continuous γ-rays in this way, and thus losing varying amounts of energy. It is evident, however, that this explanation cannot be used in the case of radium E, since the fact that one quantum only is emitted in about thirty disintegrations indicates that only a small fraction of the β-particles can lose energy in this way.

* *Proc. Roy. Soc.* 85, p. 131, 1911.

Digitized by Google

Another point of interest is that the β-ray spectrum of radium E is of about the same form and extent as most of the continuous β-ray spectra, and we should therefore expect that in these other radioactive disintegrations there is produced continuous γ-radiation of the same order of magnitude. Kovarik has shown that for radium B and radium C the emission of γ-rays is of the order of one quantum per disintegration, while, as shown above, the continuous γ-radiation gives one quantum in only 2 or 3 per cent. of the disintegrations. In the case of radium B the mean quantum is of the same order as that found for the continuous spectrum, while for radium C and most γ-ray bodies it is considerably higher. The energy of the continuous spectrum is therefore in general of the order of 1 per cent. or less of the total γ-ray energy. If this deduction is correct, the conclusion of Ahmad* that the continuous γ-radiation contains a considerable fraction of the total γ-ray emission cannot be correct. He assumed that the ordinary X-ray absorption law can be extended to the high frequency γ-rays, and it is possible that the disagreement between his work and the present conclusions is due to this assumption being untrue.

It will be seen that the conclusions of general importance for the β-ray disintegration which are reached in this paper depend on the smallness of the γ-ray effect from radium E. It seems possible both from the work of Gray and of Chadwick that a fraction of the order of 25 per cent. may originate in the screens due to the impact of the β-particles. It is difficult to be sure of this point, but should it be correct it will only tend to strengthen the deductions that have been made and by attributing all the γ-radiation to excitation within the radioactive atom we are certain to obtain an upper limit to the amount of continuous γ-rays in the ordinary β-ray disintegration and the extra amount of inhomogeneity introduced into the disintegration electrons.

I wish to express my thanks to Sir Ernest Rutherford for his interest in this work, and to Dr C. D. Ellis for helpful criticism and advice. I am also indebted to Dr Ellis and Mr Wooster for use of their measurements in finding the strength of the radium E source.

* *Proc. Roy. Soc.* 109, p. 206, 1925.

Digitized by Google

The specific heat of water vapour and the theory of the dissociation of water vapour at high temperatures. By W. H. M^cCREA, B.A., Trinity College.

[*Received* 19 August, *read* 24 October 1927.]

1. The following is an attempt to account theoretically for the observed specific heat of water vapour. The usual theory of the vibrational specific heat together with the theory of the dissociation of the H_2O molecule at high temperatures is found to account satisfactorily for the observed facts. It is shown, also, that at temperatures near 100° C. there is an effect due to polymerisation.

The experimental values of the specific heat at constant volume (C_v) at a pressure of one atmosphere are given by Partington and Shilling*. Some of them are given here in Table II and plotted in the figure.

2. The first point that arises is to explain the minimum appearing near the beginning of the curve.

In working out specific heats theoretically the molecules are always assumed to be independent systems and so, for purposes of comparison, observed values must be reduced to the ideal state of infinitely low pressure. The equation

$$\left(\frac{\partial C_v}{\partial v}\right)_T = T\left(\frac{\partial^2 p}{\partial T^2}\right)_v \quad \ldots\ldots\ldots\ldots\ldots\ldots(1)$$

is used. The effect is, of course, zero if the perfect gas equation is obeyed, but for steam the equation of state which fits the facts best is Callendar's, viz.

$$v = \frac{RT}{p} - \frac{C}{T^n} + b \ldots\ldots\ldots\ldots\ldots\ldots(2),$$

which, using (1), leads to

$$C_v(p) - C_v(0) = \frac{n(n-1)CR}{T^n v} \quad \ldots\ldots\ldots\ldots\ldots\ldots(3),$$

and here $n = \frac{10}{3}$,

$C = 1{\cdot}769 \times 10^8$ litres/gm. mol.

The calculated value is given in Table II for the first three temperatures, after which it is inappreciable. The "C" term in (2) is due to the association of the water vapour molecules and this correction to the apparent specific heat measures the amount of heat spent in separating them.

* *The Specific Heats of Gases* (Benn, 1925), Table C.

Digitized by Google

3. It is necessary to remark here that in Shilling's* more recent work the minimum is more pronounced and occurs at about 600° abs. It cannot be explained by the above method. If the values are correct the only possible explanation would seem to lie in some peculiar behaviour of the rotational specific heat. It is well known that certain expressions for the latter given by the Quantum Theory predict a maximum above the equipartition value. It is clear that such an expression added to one for the vibrational heat could yield a minimum in the total specific heat. For any measurable values of the molecular moments of inertia this could, however, only happen at about 100° abs. or lower, unless the rotational heat approached its asymptotic (equipartition) value from below by passing through a minimum after the maximum. Even then it is extremely unlikely that it would differ appreciably from equipartition at temperatures as high as 600° abs. But a number of calculations were made on different expressions given by the Quantum Mechanics and none (even allowing for the symmetry of the molecule) showed such a behaviour. It therefore seems likely that this minimum is spurious and is probably due to some difficulty in the velocity of sound method of finding specific heats.

4. Two of the fundamental frequencies of vibration of the H_2O molecule are known from band spectra data to be

$$\nu_1 = 4{\cdot}785 \times 10^{13} \text{ and } \nu_2 = 11{\cdot}235 \times 10^{13}.$$

The third (which must exist) is not definitely known, but it must be high and Dennison has suggested for it

$$\nu_3 = 21{\cdot}720 \times 10^{13}.$$

This gives for the vibrational specific heat, in the usual way,

$$C_{vib}/R = \phi\left(\frac{2283}{T}\right) + \phi\left(\frac{5362}{T}\right) + \phi\left(\frac{10360}{T}\right) \quad \ldots\ldots(4),$$

where

$$\phi(x) = x^2 \frac{e^x}{(e^x - 1)^2}.$$

The values of the specific heat from (4), assuming that the rotations have attained equipartition, are given in Table II and shown by the broken line in the graph.

There is obviously no agreement between this simple theory and experiment.

Clearly we could not have expected this theory to be sufficient since it predicts a specific heat rising asymptotically to

$$6R (= 12 \text{ approx.})$$

* *Phil. Mag.* (7), 3, p. 273 (Feb. 1927).

Digitized by Google

and the experimental curve rises well above this. Now there is known to be a certain amount of dissociation in steam at high temperatures and so some of the heat will be spent in bringing it about. It may mean, however, that there is a large general departure from the simple harmonic type of vibration since this is certain to be the case with those molecules that are just about to split up.

Before discussing the dissociation itself it is, therefore, interesting to see if this latter phenomenon is likely in itself to cause the specific heat to increase above its equipartition value. The only way of doing so seemed to be by considering the phenomenon in the case of hydrogen. Kemble and van Vleck* have given an expression which takes account of the interaction of rotation and vibration and of the departure from simple harmonic type for the H_2 molecule. I find that for a temperature of 2537° abs., which is higher than those for which they give numerical results, their formula yields a specific heat of $3{\cdot}54R$ (equipartition value $\frac{7}{2}R$; value on elementary theory $3{\cdot}03R$). Thus a law of force between the nuclei in the molecule, which gives good agreement with observation, is, in fact, capable of accounting for an excess above equipartition values.

In the figure the curves given by the elementary theory and Kemble and van Vleck's theory for H_2 are drawn for comparison with H_2O results.

5. We now consider the effect of dissociation on the specific heat. We cannot treat it by any reduction to an ideal state because the only ideal state would be that in which all the molecules were dissociated and we should then obtain the specific heat of a mixture of 1 gram molecule of H_2 and $\frac{1}{2}$ gram molecule of O_2, or even rather of 2 gram atoms of H-atoms and one of O-atoms. The only reason that we could get results this way before was that our equation of state assumed there to be no dissociation and so could be used to find the specific heat supposing the molecules removed from each other's range of influence but undissociated. As soon as the equation of state allows for dissociation it must be such as to give complete dissociation for zero pressure.

Thus the only method is to calculate the amount of heat used in splitting up the molecules at the pressure at which the specific heat is measured. This could be done empirically from experimental dissociation curves, but it is more instructive to work out the theory in the following manner, which at the same time checks the observed values.

We shall assume that, apart from dissociation, the assembly behaves as a perfect gas.

* *Phys. Rev.* xxi, p. 653 (1923).

Digitized by Google

The Dissociation of H_2O.

Suppose that the whole assembly we are considering is formed out of X atoms of oxygen and $2X$ atoms of hydrogen, and suppose that, at any instant, it consists of

M_1 free atoms of oxygen (O),
M_2 free atoms of hydrogen (H),
N_1 molecules of oxygen (O_2),
N_2 molecules of hydrogen (H_2),
N_3 molecules of water vapour (H_2O).

Then we must have

$$\left.\begin{aligned} M_1 + 2N_1 + N_3 &= X \\ M_2 + 2N_2 + 2N_3 &= 2X \end{aligned}\right\} \quad \text{.....(1).}$$

We neglect changes in the internal energies of the atoms and take as the zero energy state of the assembly that in which all the molecules are dissociated and all the atoms are at rest. The zero state of any molecule is taken to be rest in its lowest quantum state, and its heat of dissociation χ the work required to resolve it from such a state into its constituent atoms at rest at infinite separation.

Let $g_1(z)$, $g_2(z)$ be the partition functions for translatory motion of the atoms; $h_1(z)$, $h_2(z)$, $h_3(z)$ those for the molecules; $b_1(z)$, $b_2(z)$, $b_3(z)$ the partition functions for the internal motions (rotations and vibrations) of the molecules; and σ_1, σ_2, σ_3 the symmetry numbers of the molecules which are each two in the present case. Then the equilibrium state at any temperature, defined by $\theta = e^{-1/kT}$, is given by*

$$\frac{\bar{N}_1}{\bar{M}_1^2} = \frac{h_1(\theta)\, b_1(\theta)}{\sigma_1 \theta^{\chi_1} [g_1(\theta)]^2} \quad \text{.....(2),}$$

$$\frac{\bar{N}_2}{\bar{M}_2^2} = \frac{h_2(\theta)\, b_2(\theta)}{\sigma_2 \theta^{\chi_2} [g_2(\theta)]^2} \quad \text{.....(3),}$$

$$\frac{\bar{N}_3}{\bar{M}_1 \bar{M}_2^2} = \frac{h_3(\theta)\, b_3(\theta)}{\sigma_3 \theta^{\chi_3} g_1(\theta) [g_2(\theta)]^2} \quad \text{.....(4).}$$

We eliminate M_1, M_2 (dropping the bars for convenience), which for finite χ_1, χ_2, χ_3 will never be zero, though from experiment they must be negligible compared with N_1, N_2, N_3 for the temperatures we are considering†. We find

$$\frac{N_1^{\frac{1}{2}} N_2}{N_3} = \sqrt{\frac{h_1 b_1}{\sigma_1 \theta^{\chi_1} g_1^2}} \cdot \frac{h_2 b_2}{\sigma_2 \theta^{\chi_2} g_2^2} \cdot \frac{\sigma_3 \theta^{\chi_3} g_1 g_2^2}{h_3 b_3}$$

$$= \frac{h_2 b_2 h_1^{\frac{1}{2}} b_1^{\frac{1}{2}}}{2^{\frac{1}{2}} h_3 b_3}\, \theta^{\chi_3 - \chi_2 - \frac{1}{2}\chi_1} \quad \text{.....(5).}$$

* R. H. Fowler, *Phil. Mag.* (6), 45, p. 1 (1923).

† I have calculated that, for the amount of dissociation found experimentally, the value of $\frac{M_2}{N_2}$ at 1700° K is very nearly $\frac{1}{1000}$.

Digitized by Google

Neglecting M_1, M_2 in equation (1) we have $N_2 = 2N_1$. Also for the most part we are concerned with less than 1 per cent. dissociation, so that it is sufficiently accurate to consider N_3 as constant in (5) and we shall take it as applying to one gram molecule of H_2O.

To a first approximation we can write $b_i(\theta) = r_i(\theta)\, v_i(\theta)$, where $i = 1$, 2 or 3, and $r_i(\theta)$, $v_i(\theta)$ are the partition functions for rotational and vibrational motions considered separately. This is the usual first approximation used in specific heat and band spectrum theory.

For the temperatures considered all rotations will have their equipartition values, so that

$$r_j(T) = \frac{8\pi^2 I_j kT}{h^2}, \quad j = 1, 2 \quad \ldots\ldots\ldots\ldots\ldots\ldots(6),$$

for the diatomic molecules with moments of inertia I_1 and I_2, and

$$r_3(T) = \frac{8\pi^2 A kT}{h^2}\sqrt{\frac{8\pi^3 CkT}{h^2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots(6')$$

for the triatomic molecule H_2O, assuming that its moments are A, A and C*. The values of the moments are not known very accurately, so we take $I_2 = 0{\cdot}4 \times 10^{-40}$, $A = 3 \times 10^{-40}$, $C = 10^{-40}$ gm. cm.^{2}, which must be approximately correct, and, since estimates of I_1 have hitherto varied greatly, we shall take it to be $\alpha \times 10^{-40}$ and use our theory as a means of finding α which we can then compare with other estimates. [We have written $r(T)$ for the function of T obtained by substituting $e^{-1/kT}$ for θ in $r(\theta)$.]

In $v_i(\theta)$ we consider each normal mode of vibration as equivalent to a Planck oscillator of the same frequency (ν, say). Each one contributes a factor

$$1 + e^{-\frac{h\nu}{kT}} + e^{-2\frac{h\nu}{kT}} + \ldots + e^{-p\frac{h\nu}{kT}} \quad \ldots\ldots\ldots\ldots\ldots\ldots(7),$$

where the "p" state is the highest that is physically possible. It is found that $h\nu/kT$, for the values of ν in each of the three molecules and for nearly the whole range of T considered, is sufficiently large to make the expression (7) nearly unity.

Also we have†

$$h_i(\theta) = \frac{(2\pi m_i)^{\frac{3}{2}} V}{h^3 (\log 1/\theta)^{\frac{3}{2}}} = \frac{(2\pi m_i kT)^{\frac{3}{2}} V}{h^3} \quad \ldots\ldots\ldots\ldots\ldots\ldots(8),$$

where m_i is the mass of a molecule and we take

$$m_1 = 32 \times 1{\cdot}65 \times 10^{-24}, \quad m_2 \cong 2 \times 1{\cdot}65 \times 10^{-24},$$

$$m_3 = 18 \times 1{\cdot}65 \times 10^{-24} \text{ gm.},$$

* There is no simple expression for the energy levels when all the moments are unequal, but the form of $r_3(T)$ must be very similar. If two of the moments of H_2O are 3 and 1×10^{-40} the other should strictly be taken as 4 or 2×10^{-40}.

† Darwin and Fowler, *Phil. Mag.* (6), 44, p. 450 (1922).

Digitized by Google

and V is the volume at 1 atmosphere, the pressure at which the dissociation and specific heats are measured. We may write

$$V = T\frac{V_0}{273},$$

where V_0 is the theoretical volume at 273° abs.

Now let $\kappa(T)$ be the percent. dissociation at temperature $T°$ abs., then

$$N_1 = \frac{1}{2}\frac{\kappa}{100}N_3.$$

Therefore
$$\frac{N_1^{\frac{1}{2}}N_2}{N_3} = \frac{2}{N_3}\left(\frac{\kappa N_3}{200}\right)^{\frac{1}{2}} \quad \text{.....(9).}$$

Substituting from (6), (6′), (8) and (9) and writing unity for each $v_i(\theta)$, we eventually obtain from equation (5)

$$\kappa = 200\left(\frac{3.8}{81}\right)^{\frac{1}{2}}\frac{\pi^{\frac{1}{2}}(1{\cdot}65\times10^{-24})^{\frac{1}{2}}k^{\frac{1}{2}}\alpha^{\frac{1}{2}}}{(2{\cdot}705\times10^{19})^{\frac{1}{2}}h}\frac{T^{\frac{1}{2}}}{T_0^{\frac{1}{2}}}e^{\frac{2}{3}\left(\frac{\chi_3-\chi_2-\frac{1}{2}\chi_1}{kT}\right)} \quad \text{.....(10),}$$

giving
$$\log_{10}\kappa = \tfrac{5}{6}\log_{10}T - \frac{b}{T} + A \quad \text{.....(11),}$$

where A and b are constants.

Partington and Shilling* give a table of values of κ which we can use successively in equation (11). By subtracting each of the resulting equations from that immediately succeeding it we get a series of values for b and can test its constancy. The results are contained in the following Table. The experimental values of κ

TABLE I.

Temp. T (abs.)	κ observed (1 atmosphere)	b	A
°			
1300	·0029	—	—
1397	·0084	8162	1·244
1480	·0185	8032	1·052
1500	·0221	8033	1·054
1561	·0368	7932	0·986
1705	·108	8069	1·074
2155	1·18	7780	0·903
2257	1·79	7814	0·919
		Mean = 7975	Mean = 1·033

* *The Specific Heats of Gases*, p. 178.

Digitized by Google

extend to higher temperatures, but the corresponding values of b do not agree with those given below. The disagreement is so irregular that one concludes that the observations must be at fault whether or not the theory is beginning to fail. But we should, in fact, expect the theory to have to be modified since $v_3(\theta)$ is beginning to differ widely from unity.

The constancy of b and A is satisfactory, but it is not sensitive to changes in the coefficient of log T in (11) which cannot, therefore, be altogether established by it. However $(\chi_3 - \chi_2 - \frac{1}{2}\chi_1)$, and hence

TABLE II.

Abs. Temp. T	C_v observed (1 atmosphere)	Correction due to polymerisation	C_v calc. on elem. theory	Heat employed in dissociation (1 atmosphere)	Sum of last two columns
°					
373	6·65	−0·24	6·13	—	—
473	6·61	−0·10	6·34	—	—
573	6·61	−0·04	6·58	—	—
673	6·65	—	6·83	—	—
773	6·74	—	7·06	—	—
873	6·88	—	7·19	—	—
1073	7·30	—	7·70	—	—
1273	7·90	—	8·08	—	—
1473	8·70	—	8·40	·093	8·49
1673	9·67	—	8·72	—	—
1873	10·84	—	9·00	1·007	10·01
2073	12·17	—	9·26	2·285	11·55
2273	12·95	—	9·51	4·398	13·91
2473	13·45	—	9·72	—	—
2573	13·57	—	—	—	—
2673	—	—	9·92	—	—

the heat of dissociation of water vapour at 0° abs., can be calculated from b. The above value of b gives for the latter quantity

$$D_0 = 54{,}760 \text{ cals. per gm. mol.},$$

and the empirical formula for the heat of formation at any temperature is *

$$D_T = 57{,}095 + 0{\cdot}249T + 0{\cdot}0_3 9975T^2 - 0{\cdot}0_6 7798T^3 \ldots (12).$$

The agreement is quite good and it is found that D_0 is fairly sensitive to changes in the above coefficient which could not therefore be very different from that derived here.

* Partington and Shilling, *loc. cit.*, p. 169. [There seems to be a misprint in their final result for D_T, 0·449T occurring in place of 0·249T.]

Digitized by Google

From the value of A we have found, and from equation (11), a comes out to be 34·53, giving for the moment of inertia of the oxygen molecule

$$34{\cdot}53 \times 10^{-40} \text{ gm. cm.}^2.$$

This is about twice the value given by Mecke*, which is probably a reliable determination by the band spectrum using the latest theories. In view of the data the agreement here is not unsatisfactory.

The contribution to the apparent specific heat due to dissociation will be

$$\frac{D_T}{100}\frac{d}{dT}\kappa(T) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(13).$$

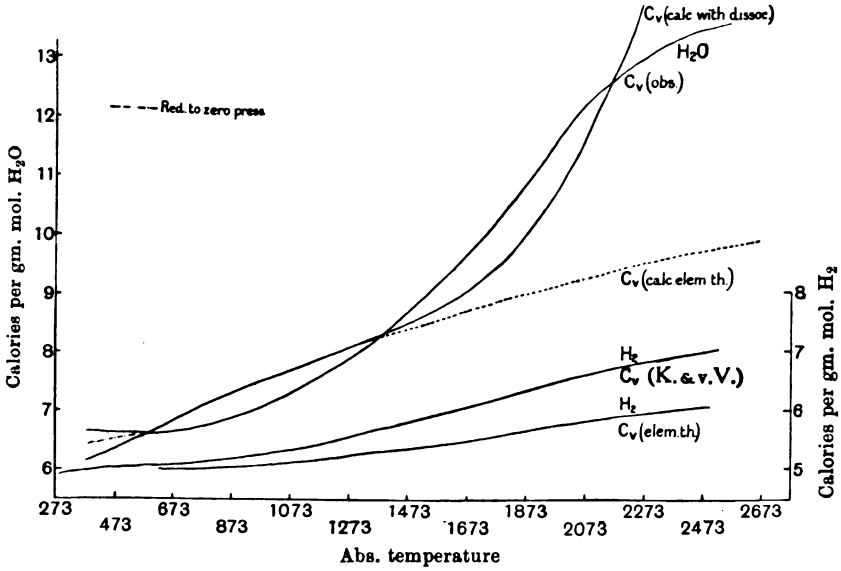


Fig. 1.

We obtain from (11) using the calculated b and A

$$\left.\begin{aligned} \log_{10}\kappa &= \tfrac{5}{6}\log_{10}T - \frac{7975}{T} + 1{\cdot}033 \\ \frac{{\cdot}4343}{\kappa}\frac{d\kappa}{dT} &= \tfrac{5}{6}({\cdot}4343)\frac{1}{T} + \frac{7975}{T^2} \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots(14).$$

The values of (13) for certain temperatures were calculated using D_T given by (12)† and the results are given in Table II and

* *Zeit. für Phys.* vol. 42, p. 396 (1927).

† D_T is derived from the observed specific heat of steam and it would be better to use the calculated specific heat based on the elementary theory since we now see that a large part of the former is due to dissociation. The variation of D_T with T is not, however, large.

Digitized by Google

shown in the figure. The agreement is good over the range in which we expect the theory to hold.

We conclude that the difference between the specific heats observed and given by the ordinary theory is due almost entirely to dissociation.

I wish to express my thanks to Mr R. H. Fowler for his interest and advice in this subject.

Digitized by Google

A Case of Double Reflexion. By E. T. S. APPLEYARD, B.A., King's College, and H. W. B. SKINNER, Ph D., Trinity College.

[*Received* 22 August, *read* 24 October 1927.]

We recently had occasion to have a right-angled quartz prism made with the optic axis of the quartz perpendicular to one of the short sides of the right-angled triangle (Fig. 1). The prism has the property that it gives two images when light enters perpendicular to one of the faces and, after internal reflexion at the hypotenuse, passes out at right angles to the other face. It was at first sight rather difficult to see why this doubling of the image occurs, since the incidence on both faces of the prism is normal and quartz is a uniaxial crystal, there can be no double refraction occurring.

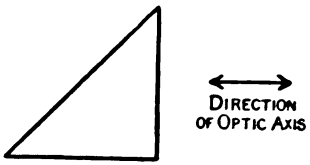


Fig. 1.

A little consideration showed that the effect was due to double reflexion at the hypotenuse of the prism. In the case of ordinary reflexion, for example from a glass surface in air, the equality of the angles of incidence and reflexion depends on the equality of the velocities of the light before and after reflexion.

In the case of the quartz prism, for the ordinary ray these two velocities are equal; for the extraordinary ray they are not, since before reflexion the light is going (for instance) parallel to the optic axis and after reflexion perpendicular to it. Thus there are two reflected rays; for one the angle of incidence is equal to the angle of reflexion, but for the other it is not. Though slightly complicated by the fact that the Huyghens wavelets are spheroids, an elementary calculation gives for the angle $\delta\phi$ between the ordinary and the extraordinary reflected rays

$$2\delta\phi = \left(\frac{V_o}{V_e}\right)^2 - 1,$$

where V_o and V_e are the principal velocities of the two waves. The angular separation of the images, calculated from the refractive constants of quartz, is 0·36°, and this agrees excellently with that actually observed. If the light enters the prism perpendicular to the optic axis, the splitting is obviously the same.

Thus the quartz prism gives an instructive example of the fact

Digitized by Google

that the angle of incidence is not always equal to the angle of reflexion, and shows on what the equality of these two angles depends. We expected to find a reference to this property in the standard works on the subject, but we have not found any explicit notice of the fact. We do not claim to have made a very elaborate search, however; for it seemed that the effect, even if it has been noticed before, is of sufficient interest to be recalled.

On account of the rotational property of the quartz, the prism behaves differently according as the light enters it parallel to the optic axis or in the perpendicular direction. If it enters parallel to the optic axis, the rotation property does not alter the character of the light internally reflected; for the rotation has no effect on natural light. In this case, looking through the prism, we see two completely polarised images. In the case when it enters by the other face of the prism, on reflexion the light is split up into two polarised components; but these components are rotated in their subsequent passage through the quartz parallel to the optic axis. Since the light entering at different points along the face of the prism traverses different thicknesses of quartz parallel to the optic axis, the effect seen, if we are using a finite aperture of the prism, is two images which are apparently more or less unpolarised. The behaviour of the prism in these two cases shows clearly that the splitting up occurs at the reflecting surface.

Thus the prism, in the second case, first splits up the light into two polarised components and then effectively depolarises it. We have found this property of the prism very useful in the spectroscopic determination of the polarisation of light. Actually we are using a spectroscope of the Littrow type, and we use the quartz prism in place of the usual single image internally reflecting prism. The spectrum lines are then doubled, the images corresponding to vibrations are parallel and perpendicular to the spectroscope slit. The advantage of this procedure is due to the depolarising effect of the prism. The light falling on the refracting prism of the spectroscope is effectively unpolarised, and hence only a very slight intensity difference between the two components occurs owing to the reflexion at an angle at the faces of the refracting prism. Intensity measurements on the two spectrum images therefore give practically the degree of polarisation directly; only a small correction has to be applied for the polarising effect of the spectroscope on light passing through it. This is a great advantage since the correction in the case of an ordinary quartz spectroscope may easily reach a value corresponding to 25 per cent. of polarisation, and owing to residual rotational effects, it varies in an extremely complicated way with the wave-length of the light. Using the present quartz reflecting prism, in the blue and ultraviolet regions, the correction is always less than 5 per cent.

Digitized by Google

A Convenient Method of Distillation of the Alkali Metals. By R. J. CLARK, Lecturer in Physics and Carnegie Teaching Fellow in the University of Edinburgh. (Communicated by Prof. Sir E. RUTHERFORD.)

[*Received* 8 August, *read* 24 October 1927.]

The alkali metals, sodium, potassium and rubidium can be distilled easily in a good vacuum and obtained reasonably free from occluded gas in the following way. As potassium is now used for the absorption of mercury vapour and many experiments are being done on the others, an account of a convenient way of preparing pure specimens may be of some service to experimenters.

A still is first made like that shown in Fig. 1, which will serve for the preparation of trebly distilled specimens. The size shown will do for the distillation of about six or seven grams. The constrictions where the various sealings are to be done should be of only moderately thick walls, and the bore should be about 1·5 to 2 mm. The connections are on one side only for convenience in washing. When the still is made it is thoroughly cleaned and dried, and set vertically in a clamp (on the tube *S* by preference) and connected through a mercury-vapour trap to either a Gaede rotatory mercury-pump or to a small diffusion pump.

A sufficient quantity of the alkali metal is washed in some dry ether containing 4 or 5 per cent. of alcohol, and all the oily scale is knocked off it. After two or three washings when it is bright and clean beneath the ether, it is taken out, wiped off with filter paper and dropped into the still. A *small* wad of glass wool is next put in and the still is sealed off at *A* and exhausted.

As soon as the exhaustion is well under way all the still except *B* is heated with a Bunsen burner (non-luminous flame) and then allowed to cool. *B* is next heated *very gently* when the metal will melt, and breaking through its skin of oxide, run into the bulb *C*. If this heating is properly done the oxide will not fuse and it will be easy to seal *B* off. The residue left behind in *B* will be contaminated with oil, and it is important that none of it should get into *C*. When the metal melts a lot of gas comes off and the pump should be watched to see that it does not choke.

A small electrically heated oven is next inverted over *C*, and the metal distilled from one bulb to the next, the operator sealing off the bulbs in turn as they empty. The open end of the oven should come a little below the top of the next lower bulb, in order to keep the connecting tube warm. Finally, the metal may be run into an ampoule which is to be sealed off, or it may be run directly into the apparatus in which it is to be used.

Digitized by Google

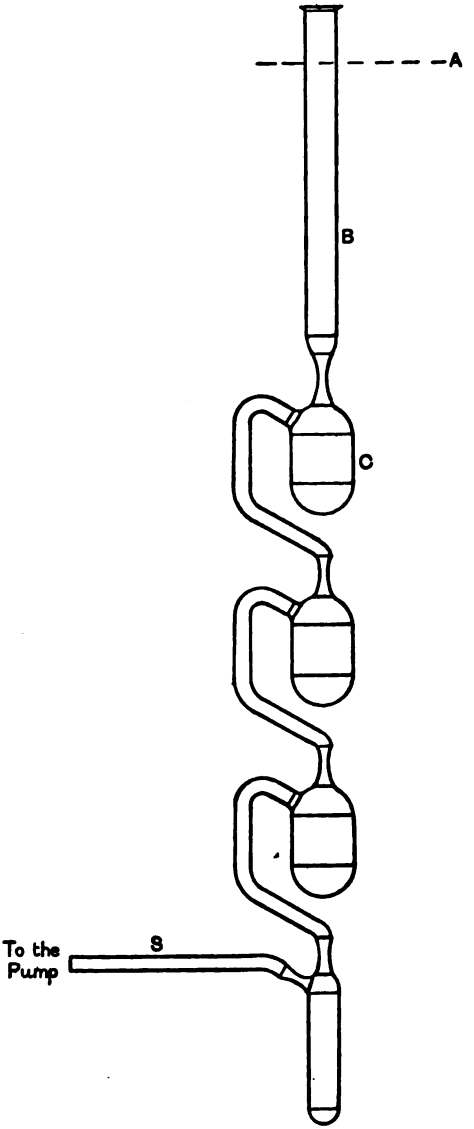


Fig. 1. (Scale $\frac{1}{3}$ natural size)

Digitized by Google

To distil sodium it is best to use either Pyrex (Jobling) or Monax (Moncrieff) or similar glasses; the temperature of the oven should be about 500 degrees centigrade. Potassium and rubidium can be distilled in soda-glass and a temperature of 400 degrees centigrade is ample for potassium and 350 for rubidium.

A suitable oven can be made from a piece of 20 gauge sheet iron pipe, 3 inches in diameter and 6 inches long. Over one layer of asbestos cloth a single winding of nichrome wire about number 22 S.W.G. is put on (the resistance will be about 25 ohms); this winding is covered with about three-quarter inch thickness of asbestos, and the whole is slipped inside a piece of 5 inch sheet iron pipe 7 inches long. One end is closed with three-quarter inch asbestos and a sheet iron disc on which the binding screws are mounted. This oven can be connected directly to the 110 volt mains. The temperature is regulated by a series rheostat and in practice the rate of distillation is a good guide to the proper temperature. Each distillation of 5 grams should take about 10 to 15 minutes.

This method has been used also to distil mercuric-chloride, mercurous-chloride, zinc bromide, arsenic trioxide and some similar substances in order to free them from occluded gases and non-volatile impurities. Ordinary soda-glass is suitable for all these substances, and a person skilled in this sort of manipulation will find it the most convenient, but any one else had better use a harder glass, as the chance of cracking a tube in sealing it, is then very much less.

Digitized by Google

A Rapid Mercury Still. By R. J. CLARK, Lecturer in Physics and Carnegie Teaching Fellow in the University of Edinburgh. (Communicated by Prof. Sir E. RUTHERFORD.)

[*Received* 8 August, *read* 24 October 1927.]

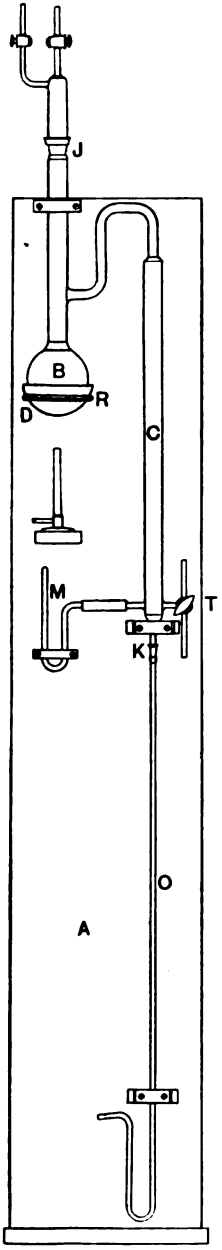
This still is a modification of Hulett's pattern* in which mercury is distilled in a current of air at greatly reduced pressure. The current of air has two uses; first it prevents bumping and so is led in beneath the surface of the boiling mercury, and secondly, it oxidizes most of the zinc, cadmium and lead which go over. The oxides float on the distillate and are easily removed by straining it. A still made according to Hulett's directions and having a boiler of half a litre capacity will distil satisfactorily about 750 grams per hour. If the rate of distillation is increased much, some of the other metals go over, and if more air is admitted to prevent this, the boiling mercury is thrown about very violently in the boiler, which breaks sooner or later.

By one change in design, the use of two inlets for air instead of one, the capacity of the still can be doubled for a given size of boiler. One of the two inlets is near the bottom of the boiler beneath the mercury, and the other is at the top above the surface. Only enough air to stop bumping is led in through the first, the rest goes through the other. A boiler of half a litre capacity can be used to distil 1500 grams per hour in this way.

A pattern of still, one of which has been in use for some years, is shown in the accompanying diagram, which is to scale and indicates sufficiently the way in which the still is made. The boiler *B*, a round-bottomed flask (soda-glass) of half a litre capacity rests on some fine brass filings in a hemispherical iron dish *D*, of 6 cm. radius, in its turn supported on an iron ring *R* fixed to the wooden support *A*. The neck of the flask is 2·5 cm. in diameter and 30 cm. long so that the mercury sealed ground joint *J* remains cool. Through this joint pass two concentric tubes, the larger extending 2 cm. into the boiler when in position, the smaller ending in a fine capillary 1 or 2 mm. only from the bottom of the boiler. This is shown in the detailed diagram of the boiler in Fig. 3. Taps are provided to regulate the amount of air admitted through each tube. The eduction tube to the condenser is 15 mm. in diameter, and is sealed in the neck 7 cm. above the bulb, and to avoid priming it is bent at right angles, and carried up 20 cm.; then it is bent through 180° on a radius of 4 cm. and joins the condenser *C*. This is a vertical tube 3 cm. in diameter, 60 cm.

* Hulett, *Physical Review*, vol. 21, p. 288, 1905.

Digitized by Google



Scale $\frac{1}{12{\cdot}5}$ natural size.

Fig. 1.

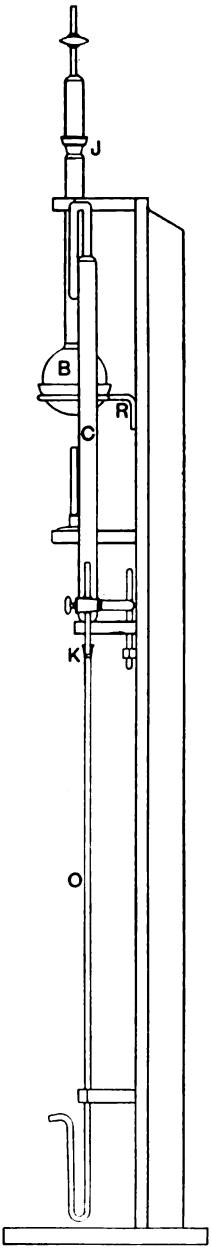
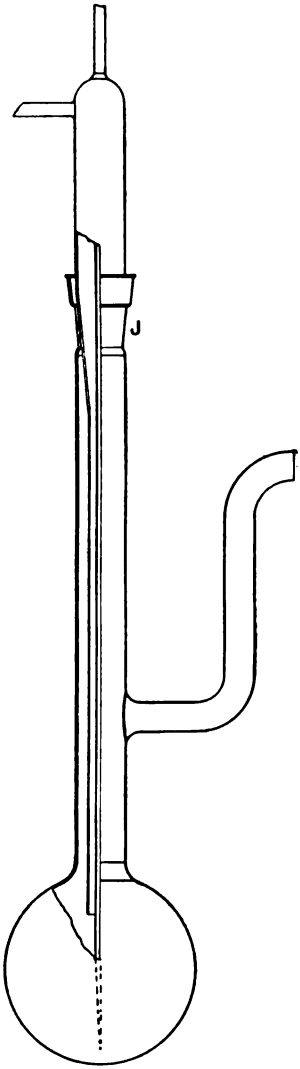


Fig. 2.



Scale $\frac{1}{4}$ (approx.) natural size.

Fig. 3.

Digitized by Google

long and with a wall ·75 mm. thick; it is supported on a wooden arm at its lower end, and through a hole in this arm passes a 7 mm. tube ending in the male part of a ground joint *K*. A three-way tap *T* is sealed on 2 cm. above the bottom of the condenser, and so the still can be evacuated or filled with air. The other part of this ground joint is at the upper end of the barometric over-flow *O*, which is a heavy wall capillary tubing about 1 mm. in bore. A small manometer is connected to the lower end of the condenser by a piece of pressure tubing. The whole apparatus is mounted on a board stiffened by ribs, and occupies a floor space 30 cm. square, and is 150 cm. high.

To work the still the upper part of the head is taken off, and 4500 grams of mercury are poured into the boiler through a funnel with a stem long enough to reach nearly to the bottom. The joint *J* is rubbed over with graphite and put in place, and a little mercury is put in the cup to seal it. A *good* water suction pump is next connected to the tap *T* by some wide tubing. It is important that there should be no leaks and that the passage of gas to the pump should be entirely unobstructed. The end of the over-flow tube is now dipped into a dish of mercury, and the tap *T* is turned to connect the still with the pump. When the manometer shows a pressure of 25 mm. the amount of air admitted by the submerged inlet is adjusted so that one or two bubbles per second rise through the mercury. The other inlet is then adjusted until the pressure stays constant at about 21 mm. The burner is then lighted; when the distillation begins the air supply will probably want further adjustment. Once started it can be let alone until there is 400 grams or so left in the boiler when the connection with the pump is cut off and the still allowed to cool. The distillate is then strained once or twice through a perforated filter-paper.

A still like this works as a pump, the pressure of oxygen being many times greater in the cold part of the condenser than in the boiler. Calculation shows that when 1500 grams of mercury are distilled per hour, 1 c.c. of air is admitted per second and the manometer shows a pressure of 20 mm., the pressure of oxygen in the boiler is less than 0·1 mm. Plenty of air must therefore be admitted and pumped out again by an efficient pump.

Digitized by Google

Notes on Spectrophotometry. By R. W. DITCHBURN, B.A., Trinity College, Isaac Newton Student.

[*Received and read* 25 July 1927.]

1. *Sources of light for ultra-violet spectrophotometry.*

The source of light used as a background is an important factor in determining the convenience and accuracy of ultra-violet absorption work, etc. If a source of light of constant intensity is available, a direct comparison method can be used and it is only necessary to calibrate the plates. If the source of light is not constant in intensity, it is necessary to divide the light into two beams and use one to check the variations of intensity while the other goes through the absorbing substance or (during calibration) the reducing sector or wedge. This latter method requires much more complicated apparatus and if the variations in the source are at all large it becomes inaccurate. In addition to being constant in intensity a good background for ultra-violet absorption spectra should possess the following qualities:

(1) Most of the energy should be emitted in the form of a continuous spectrum.

(2) It is desirable to be able to use one photograph of the whole region to be investigated. For this purpose it is necessary that the variations of intensity in different parts of the spectrum should be small enough for it to be possible to arrange the exposure so that all parts of the spectrum are within the correct exposure range, i.e. it must not be necessary to over-expose any part in order to get a strong enough intensity at another wave-length.

The hydrogen continuous spectrum possesses both these qualities and is an excellent background for the region on the short wave-length side of 3200 A.U. It may be used for longer wave-lengths, but the hydrogen secondary lines are apt to prove troublesome unless a fairly large dispersion is used.

The most convenient way of producing this light is a discharge tube of the pattern invented by Lyman* (Fig. 1). The writer has found that with suitable precautions it is possible to make this tube run perfectly steadily. In the form used by the writer the tube is of soft glass and a quartz window is waxed on to one end, enabling the capillary to be viewed end on through the cylindrical electrode. A more permanent tube might be constructed of pyrex with tungsten-pyrex electrode seals and a pyrex-quartz seal enabling the quartz window to be fused on.

* *Astrophysical Journ.* 23, 181 (1906).

Digitized by Google

The diameter of the capillary should be between 2 and 3 mms. and the length between 5 and 6 cms. There is some advantage to be gained by making one of the ends larger than the other and, since the window must be placed well away from the electrode, the window end should be the larger. A bulb of about a litre capacity is attached to the tube to prevent violent changes of pressure when the tube heats up and also to reduce the fall of pressure produced by prolonged running of the discharge. The best pressure is about a millimetre of mercury, which is a little above the pressure at which the effective impedance of the tube is least. This assists in producing stability. When the current increases, the tube becomes hotter, the pressure rises and causes the current to fall, and similarly if the current falls the consequent change of pressure tends to bring it back to its former value.

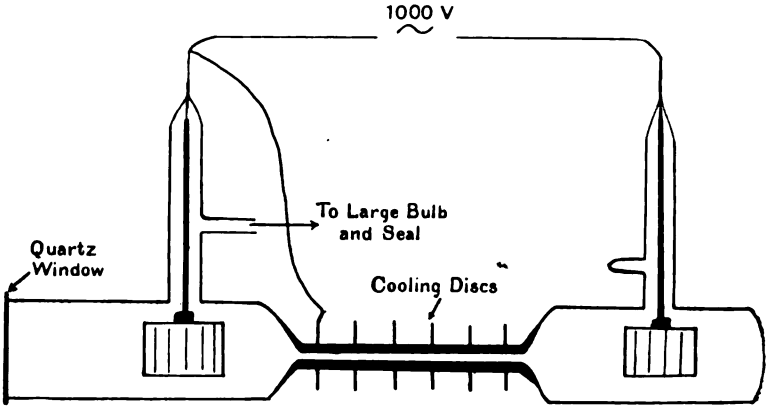


Fig. 1.

It is very desirable to have the hydrogen really pure. For this purpose the tube should be run for some hours with a stream of electrolytic hydrogen passing through it, liquid air traps being placed between it and the nearest taps. This removes mercury and hydro-carbons. Lyman* found that when the discharge was run a film collected on the window. This seems to be entirely due to the presence of hydro-carbons. For this reason the window should be taken off and cleaned after the discharge has been run long enough to remove all traces of hydro-carbon. After this the tube is again run for some hours with the stream passing and then the pressure is adjusted to a little above the right value and the taps are turned off. The tube is then run for at least three hours more to get rid of the last traces of oxygen and mercury and is then sealed off. The purity

* Lyman, *loc. cit.*

Digitized by Google

of the hydrogen is best tested by watching the relative intensity of the line and continuous spectra. When the hydrogen becomes really pure the relative intensity of the latter increases enormously.

The tube runs much more steadily with nickel electrodes than with aluminium. The electrodes should be prevented from vibrating and this is conveniently done with small glass wedges. The source of electrical supply is a 1000 volt transformer, the current being controlled by a resistance in the primary. As much as 200 milliampères may be put through the tube steadily, but with currents of more than 30 milliampères the capillary needs cooling. This is secured by wrapping with tinfoil on which may be fastened the brass radiating discs. With a current of 100 milliampères the continuous spectrum is sufficiently intense to be photographed in a low dispersion spectrograph with an exposure of a few seconds.

The tube may still run unsteadily owing to certain capacity effects. These are avoided in the case of the tube at present in use by connecting the cooler to one side of the supply. For another tube it was found to be desirable to connect a Leyden jar to the cooler instead. These capacity effects are difficult to understand, but it is fairly easy to find the right arrangement to eliminate them after a few trials.

When this has been done the tube runs steadily to better than a half per cent. If run on the town supply it is of course necessary to follow the fluctuations of the supply and keep the current through the tube constant by adjusting the resistance in the primary of the transformer. Photometric tests showed that for small changes the intensity of light produced is proportional to the current. Thus the constancy of the source is simply dependent on the effectiveness of the control of the supply voltage.

The only other constant sources of continuous ultra-violet light are the tungsten lamps and the cadmium spark used by Harrison*. The intensity of the tungsten lamp falls very rapidly in the direction of shorter wave-length. It is necessary to take several exposures to cover any considerable range of wave-length and the intensity is very feeble for wave-lengths below 2800 A.U. The source used by Harrison seems to possess most of the qualities of the hydrogen tube in the region from 3500 A.U. to 2000 A.U. The hydrogen discharge tube appears to be simpler and more convenient than the spark. Moreover, with a fluorite window, it may be used for work in the Lyman region.

2. *Reducing Sectors, etc.*

For all methods of spectrophotometry it is necessary to have some way of reducing a beam of light in a known ratio. The purpose of this section is to discuss the various devices which have

* Harrison, *Phys. Rev.* 24, 466 (1924).

Digitized by Google

been used, and to describe a form of sector which is simple in construction and possesses many advantages.

I. *Neutral Wedge Methods**. A wedge of neutral grey glass is cemented to one of clear glass. The difference between the absorption coefficient (for a given wavelength) at two planes, such as *A* and *B*, is proportional to the distance between them. If it is desired to have a small area where the absorption coefficient is constant, two grey wedges may be used, as shown in Fig. 2 *b*. The small wedge has the same angle as the other, and thus the same density gradient. The density of the combination is constant over the area of the small wedge, and is varied by moving the large wedge.

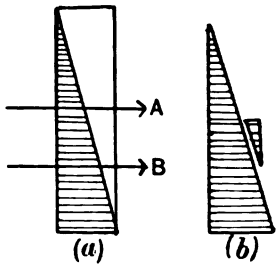


Fig. 2.

For ultra-violet work, wedges of gelatine on quartz have been investigated to λ2900 A.U. by Toy and Ghosh†. Wedges of sputtered platinum have been used by Merton‡ and by Dorgelo§. These last have not a uniform density gradient, but this can be allowed for in calibration.

The neutral wedge is very useful for many purposes (as is illustrated in the book of Dobson, Griffiths and Harrison), and it is the basis of the very interesting method of spectrophotometry devised by Merton, but it is at best a secondary standard in that it needs to be calibrated by reference to some other reducing device.

II. *Variable Aperture Screens.* The screen is used to reduce the effective area of one of the lenses in the optical system. If the illumination were uniform over the whole area of the lens an iris diaphragm might be used. To compensate for uneven illumination a grating or gauze replaces the simpler diaphragm. The opaque and clear spaces must be quite irregularly distributed, or else the screen must be kept in irregular motion to and fro across the surface of the lens.

The calibration is made by measuring up the area of the opaque spaces. A set of about five different screens is needed, and by using these in different combinations a sufficient number of calibration points is obtained.

This method of reducing the intensity suffers from two disadvantages.

* Dobson, Griffiths and Harrison, *Photometric Photometry*, ch. I.
† Toy and Ghosh, *Phil. Mag.* 40, 775 (1920).
‡ Merton, *Proc. Roy. Soc.* A, 106, 378 (1924).
§ Dorgelo, *Phys. Zeits.* 26, 756 (1925); *Zeits. f. Phys.* 31, 827 (1925).

Digitized by Google

(1) If the size of the spaces is too small diffraction effects become appreciable, and these cause an error which is different for different wave-lengths. If, on the other hand, the size is too large the distribution over the lens area will not be sufficiently irregular. It is very difficult to be sure that both these troubles have been avoided.

(2) The calibration (which must consist of measuring up each space) is rather a lengthy process, and it is by no means certain that the value will remain constant. If the screen consists of a set of lines ruled on a silvered surface it is very liable to collect a little dust or to become oxidised, and so change its opacity by a few per cent. Wire gauze or metal plates drilled with a number of holes are a good deal better from this point of view, but not perfectly satisfactory.

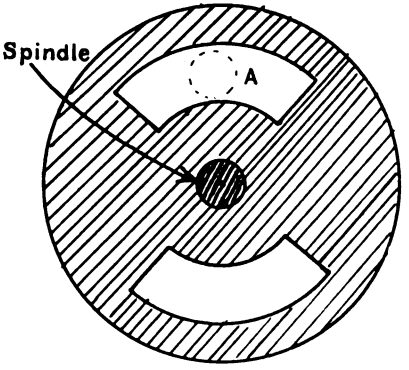


Fig. 3.

III. *Nicol Prisms.* A pair of nicol prisms is placed in the path of the beam; the intensity of the light transmitted is proportional to the square of the cosine of the angle between their principal planes. It is necessary to arrange that the light entering the spectroscope always has the same plane of polarisation.

This method is convenient for visual measurements though, as will be shown later, the angle-sensitivity is rather low. For ultra-violet work Foucault air prisms might be used since Canada Balsam is opaque. These have a small field, and cannot be used below $\lambda 2500$ A.U., since calcite begins to absorb strongly at this wave-length.

IV. *The Circumference-aperture Sector.* In this device two discs of the shape shown in Fig. 3 are fastened on the same spindle so that they can be fixed at any angle relative to each other and so that the whole system can be rotated. The light passes through the aperture A, of which the size can be varied by

Digitized by Google

altering the angle between the discs. This device is very simple and is quite useful for finding the shape of an absorption curve (qualitatively), but it cannot be used for accurate quantitative work on account of the failure of the reciprocity law of photographic action. It is known that the effect of steady exposures is approximately proportional to It^p (where I is the intensity, t the time of exposure and p is a constant). These, however, are not steady exposures, and it is not known what the law of action is for intermittent exposures of this type. There is thus no possibility of working out a correction, and this method is more uncertain than calibrating by reducing the length of exposure.

V. *The Judd-Lewis Vane Sector**. The sector system consists of four triangular vanes, each of which has two edges at right angles. The angles of the vanes all meet in a point, so that when they are all in one plane (perpendicular to the direction of the beam) no light is allowed to pass. The vanes are capable of being rotated about their bisectors ("m" in Fig. 4 *a*) through a measured angle, and their shadow in the beam is in the form of a maltese cross (Fig. 4 *b*).

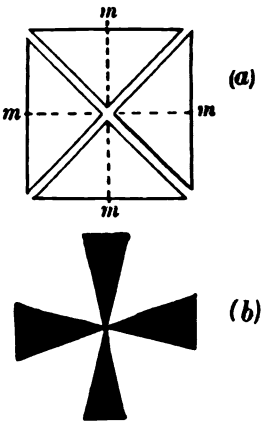


Fig. 4.

The sector system is placed in a parallel beam of light with its centre as near as possible to the axis of the optical system. The area of the beam is thus reduced in a known ratio. The intensity of the beam will be reduced in the same ratio, provided that the mean intensity over the shadowed part of the beam is equal to that over the clear part. In view of the form of the shadow this condition will in general be fulfilled, but if the source is an arc or spark capable of wandering far from the axis of the optical system an appreciable error might be involved. The sector possesses the advantage of having the whole aperture available (and not half, as in most sectors). This advantage is however completely outweighed by the mechanical difficulties of the device. These make it useless for general purposes.

VI. *The Radial-aperture Sector.*

We shall now describe a radial-aperture sector which was made for the writer by Messrs Bellingham and Stanley, and has been in use for the past year.

* Judd-Lewis, *Trans. Chem. Soc.* 1919, 312; *Journ. Roy. Soc. of Arts*, 1921, 206.

Digitized by Google

This sector consists essentially of two discs A and B, in each of which there is a semicircular hole. The discs have a common centre C, being mounted on the same large hollow spindle D. A can be rotated relative to B and fixed in any position. A is divided into degrees, and B carries a vernier so that the angle LCM can be read to a tenth of a degree. The sector system is placed in a parallel beam of light, and C is fixed as accurately as possible on the axis of the optical system. The reduction of area and of intensity is similar to that produced by the Judd-Lewis sector, but the radial aperture sector may be rotated to eliminate the effects of non-uniformity over the area of the beam. It has always been found necessary to carry out this rotation. When the centre of the

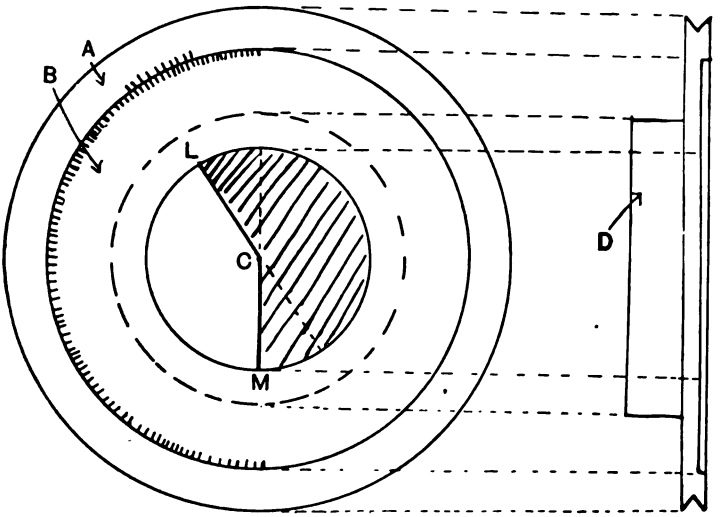


Fig. 5.

sector has apparently been set on the optical axis, it is found on looking into the spectroscope that quite large changes of intensity are produced when the sector is turned. Great care is needed in making a final adjustment so that no perceptible difference could be seen on rotating the sector.

Dr Skinner has pointed out to me that the difficulty of this adjustment may be avoided when it is possible to place the sector in one focal plane of the lens which condenses the light on to the slit of the spectrograph. When the sector is not quite accurately set on the axis of the optical system, intermittency is really produced by movement of the image of the sector which, in general, is partially focussed on the slit of the spectrograph. With the above arrangement, however, the image of the sector is at infinity

Digitized by Google

and is thus completely out of focus on the slit of the spectrograph. It is desirable to employ this arrangement when using any of the aperture reducing devices described in Sections II and IV.

In comparison with the Judd-Lewis instrument this sector has the disadvantages that only one half of the field is available, and that it has only one aperture instead of four symmetrically placed openings. It has, on the other hand, the advantage of being much simpler in construction (which means that less reliance is placed on the workmanship and that it is less costly). Moreover it may be rotated when this is desirable.

For technical purposes and routine work it may be desirable to have an instrument which need not be rotated. For this purpose it is necessary to have several symmetrical openings. The radial aperture sector may be modified in the following way. Two metal plates are accurately cut in the shape shown in the diagram and mounted on a large hollow spindle. By fixing them at different angles relative to one another the beam may be cut down in any desired ratio. In this way three symmetrically placed openings are obtained. By suitably cutting the plates any number of symmetrically placed openings may be obtained. If more openings than one are used, however, each of them will have to be proportionately narrower for a given intensity. This means that diffraction effects will be greater and also (as will be shown later) that the angle-sensitivity will be reduced in proportion to the number of openings. For purposes where it is not desired to rotate the sector, the three-aperture type is probably the best. It should be noted that even this type, while not so simple in construction as the single-aperture type, is considerably simpler than the vane sector and can be rotated if necessary.

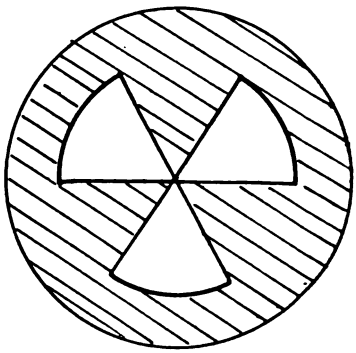


Fig. 6.

Angle Sensitivities.

In several of the methods described above the reduced intensity is calculated from the measured value of an angle. In comparing the instruments it is interesting to see how the intensity changes with the angle in each case. We may define S by the following equation:

$$\frac{1}{S} = \frac{1}{I}\frac{dI}{d\theta} \quad \text{.....(1),}$$

where I is the fraction of the maximum intensity which is trans-

Digitized by Google

mitted when the measured angle is θ (zero angle being that where no light is transmitted). S measures the accuracy with which it is possible to measure the absorption coefficient of a substance which transmits the same fraction of the incident light. Or, looking at it from a slightly different point of view, S measures the rate of change of density of an equivalent absorbing screen; for

$$D = \log_{10} \frac{1}{I} = k \log_e I$$

$$\left(\text{where } k = \frac{-1}{2{\cdot}3}\right),$$

$$\frac{1}{S} = \frac{d}{d\theta}(\log I)$$

$$= k \frac{d}{d\theta}(D) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2).$$

We shall refer to S as the "angle-sensitivity," and proceed to compare its values for the different instruments.

We first write down the relations between I and θ:

(A) *For the Nicol prisms,*

$$I = \sin^2 \theta \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3);$$

(B) *For the Judd-Lewis vane sector**,

$$I = \frac{\pi - 4 \tan^{-1}(\cos \theta)}{\pi} \quad \ldots\ldots\ldots\ldots\ldots\ldots (4);$$

(C) *For the Radial-aperture sector,*

$$I = \frac{n\theta}{\pi} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (5),$$

where n is the number of openings available.

For the circumference-aperture sector the relation is the same as for the radial-aperture sector with two openings.

From these we obtain for S,

(A) *For the Nicol prisms,*

$$S = \tfrac{1}{2} \tan \theta,$$

$$S = \tfrac{1}{2} \sqrt{\frac{I}{1 + I}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (6);$$

(C) *For the Radial-aperture sector,*

$$S = \theta,$$

$$S = \frac{\pi I}{n} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (7).$$

(Thus S is inversely proportional to n, the number of openings.)

* Judd-Lewis, *Trans. Chem. Soc.* 1919, 316.

Digitized by Google

(B) The expression for S is rather complicated in the case of the vane sector, and the values of S are best obtained by calculating pairs of neighbouring values for I.

The values of S for the different instruments are shown as functions of I in the accompanying graph (Fig. 7). The graph shows that, for intensities of ten per cent. and less, the angle-sensitivities of the Judd-Lewis and the radial-aperture sectors (with

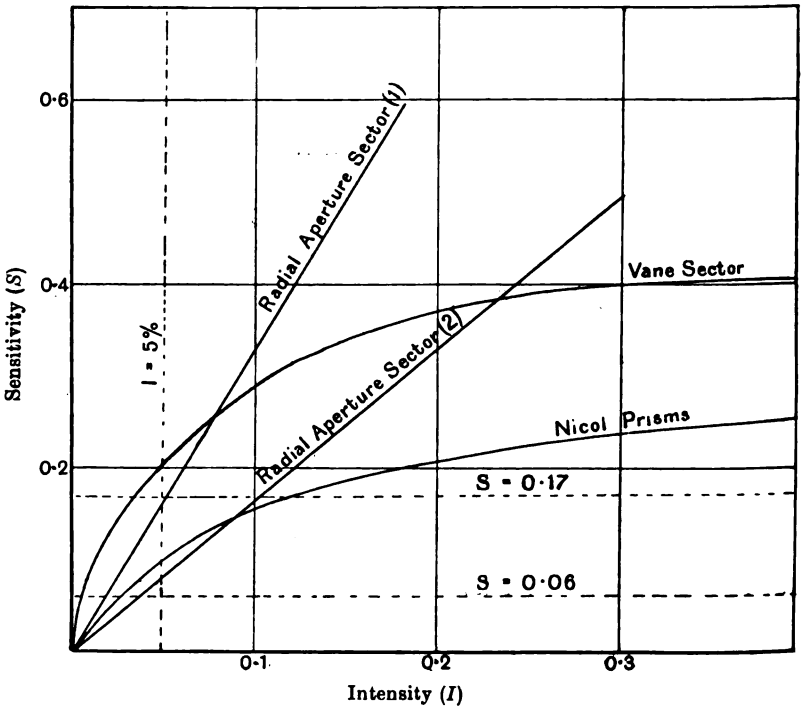


Fig. 7.

one opening) are about equal. The angle-sensitivity of the nicols in this region is about the same as that of the radial-aperture sector with two openings, i.e. half that of the radial-aperture sector with one opening. For intensities of over ten per cent. the radial-aperture sector with one opening has a far greater angle-sensitivity than any other device.

In general it is necessary to be able to measure I to one per cent. in order to have the error of the sector of the same order as other experimental errors.

Digitized by Google

Since $S = Id\theta/dI$ and we must have $dI/I \nless 1/100$, we require $S \nless \cdot 17$ if angle readings are correct to one-tenth of a degree, and $S \nless \cdot 06$ if angle readings are correct to one-thirtieth of a degree.

From the graph we have

$$S \nless \cdot 17$$

for the	Vane sector when	$I \nless \cdot 04$
„	„ Radial-aperture sector (1 opening)	$I \nless \cdot 06$
„	„ Radial-aperture sector (2 openings)	$I \nless \cdot 11$
„	„ Nicols	$I \nless \cdot 12$

and $S \nless \cdot 06$

for the	Vane sector when	$I \nless \cdot 007$
„	„ Radial-aperture sector (1 opening)	$I \nless \cdot 02$
„	„ Radial-aperture sector (2 openings)	$I \nless \cdot 04$
„	„ Nicols	$I \nless \cdot 03$.

In general it is not desirable to use this type of device to effect a reduction to less than about five or ten per cent. of the original intensity. In the case of the sectors, diffraction effects become appreciable when the openings are too narrow. In the case of the nicols, errors due to obliquity of part of the light may enter. If a greater reduction is desired it is better to use a wedge and calibrate it in position by reference to the sector or the nicols. We thus see that an accuracy of one per cent. in I is required for values of I of ·05 and greater. For this purpose it is necessary to know θ to one-tenth of a degree for the vane sector and the radial-aperture sector (with one opening). For the nicols and the radial-aperture sectors with two or three openings it is necessary to read the angle to one-thirtieth of a degree in order to reach this accuracy.

If it is desired to be able to vary I by one per cent. it is necessary to have verniers reading to one-tenth or one-thirtieth of a degree. But if, as is usual, it is only necessary to know I to this degree of accuracy for a set of values spaced at about ten per cent. intervals, the scale need only be divided into degrees, provided that the setting is made with the above accuracy.

Digitized by Google

The Absorption and Reduction in Velocity of β-rays on their passage through matter. By E. MADGWICK, M.C., Ph.D., Emmanuel College. (Communicated by Prof. Sir ERNEST RUTHERFORD.)

[*Received* 22 August, *read* 24 October 1927.]

Introduction. The early experiments of Lenard (1), Leithauser (2), and Becker (3) on cathode rays led to the exponential law of absorption; those of Rutherford (4), Schmidt (5) and Hahn and Meitner (6) on β-rays indicated that the rays from some radioactive bodies were absorbed exponentially and that those from others were not. It was considered by Hahn and Meitner that the exponential law of absorption was an indication of the homogeneity of the rays. It was found also that the exponential law was not disturbed by the thickness of radiating material and this led to the conclusion that the rays are not appreciably reduced in velocity on their passage through matter. The theories of McClelland (7) and Schmidt (8) attributed the absorption to the stoppage of particles in mid-career.

In 1906 Sir J. J. Thomson advanced a theory which supposes the energy lost by charged particles on their passage through matter to be transferred to the constituents of the atoms in the course of a series of collisions. Suppose a colliding particle of charge E, velocity V and mass M to be projected towards an electron of charge ρ and mass m in a direction such that p is the perpendicular distance from the electron to the original path of the colliding particle. Then if the electron may be regarded as free the energy transferred is

$$\frac{2E^2e^2}{mV^2}\frac{1}{p^2+a^2},$$

where

$$a=\frac{eE(M+m)}{MmV^2}.$$

If N is the number of atoms per unit volume and n the number of electrons per atom the average loss of kinetic energy in penetrating a distance Δx is

$$\Delta T=\frac{4\pi e^2E^2Nn\Delta x}{mV^2}\int_0^p\frac{pdp}{p^2+a^2} \quad \ldots\ldots\ldots\ldots\ldots(1),$$

the effect of the nuclei on decrease in energy being negligibly small. If the integration be carried out between $p=0$ and $p=\infty$, the value of ΔT is infinite, and Thomson fixed an upper limit by considering only collisions in which the closest distance of approach is of the order of the average distance apart of the electrons

Digitized by Google

within the atom. When $\frac{1}{2}MV^2$ may be written for T, equation (1) gives by integration

$$V_0^4 - V_x^4 = Cx \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2),$$

where C is constant. Experimenting with homogeneous cathode rays Whiddington (9) verified equation (2) and found C for various elements. Wilson (10), using the electrical method, showed that homogeneous β-rays are absorbed by aluminium according to the linear law

$$I = k(a - x) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3),$$

and that the exponential law previously found was in fact a consequence of the inhomogeneity of the radiation employed. He also measured the reduction in velocity of homogeneous β-rays on their passage through sheets of various thicknesses of absorbing material. Measurements of the reduction in velocity of β-rays have been carried out photographically by Danysz (11), von Baeyer (12), Rawlinson (13) and d'Espine (14).

Bohr's theory. Bohr (15) decided on an upper limit for the integration in equation (1) by assuming, as in the electron theory of dispersion, that the electrons when disturbed vibrate with characteristic frequencies ν. He introduced the conception of a time of collision, i.e. a time interval of the same order as that which the colliding particle takes to travel a distance p. If the time of collision is very short compared with the time of vibration of the electron the energy transfer will be nearly the same as if the electron were free; if this time is long compared with the time of vibration the electron will behave almost as if rigidly bound. From these considerations Bohr found for the effective upper limit of the integral $\frac{kV}{2\pi\nu}$, where k has the value 1·123, deduced the equation (2), and calculated the constant C. In addition to the average loss of energy he estimated the most probable loss which is suffered by the individual particles and also allowed for the fact that V may be comparable with c, the velocity of light. For β-rays, remembering that $M = m$ and $E = e$, his expression is

$$\Delta T = \frac{2\pi e^4 N \Delta x}{mV^2} \sum_{r=1}^{n} \left[\log \left(\frac{k^2 V^2 N n \Delta x}{4\pi \nu_r^2} \right) - \log (1 - \beta^2) - \beta^2 \right] \quad (\beta = V/c) \quad \ldots\ldots(4),$$

the electrons in the atom having characteristic frequencies ν_1, ν_2, ... ν_n. If we suppose this law to hold until the particles are stopped we have for the "range" of the β-rays,

$$R = \int_0^R \Delta x = \int_0^T \frac{mc^2\beta^2}{2\pi e^4 N \Sigma} \Delta T,$$

Digitized by Google

writing Σ for the last factor in equation (4). Considering this as constant

$$R = \frac{m^2c^4}{2\pi e^4 N\Sigma}\int_0^\beta \frac{\beta^3 d\beta}{(1-\beta^2)^{\frac{5}{2}}}$$

$$= \frac{m^2c^4}{2\pi e^4 N\Sigma}\{(1-\beta^2)^{\frac{1}{2}} + (1-\beta^2)^{-\frac{1}{2}} - 2\} \quad \ldots\ldots\ldots(5).$$

Bohr's theory can thus be tested by measurements both of the reduction in velocity and of the ranges of homogeneous β-rays. In the case of α-rays the theory gives an excellent agreement with the velocity-distance curve in the case of elements of low atomic number (for which certain approximations introduced by Bohr are valid) and from experiments on aluminium the value 39·02 is found for $\frac{1}{n}\Sigma \log \nu$. This is employed in estimating Σ in the above equations. It is convenient to express the results of experiments on reduction in velocity in terms of $\Delta(H\rho)$. Since

$$H\rho = \frac{c^2 m}{e}\beta(1-\beta^2)^{-\frac{1}{2}},$$

and

$$T = c^2 m\{(1-\beta^2)^{-\frac{1}{2}} - 1\},$$

we have $\Delta T = e\beta\Delta(H\rho)$ and, from equation (4),

$$\Delta(H\rho) = \frac{2\pi e^3 N\Delta x}{mc^2\beta^3}\sum_{r=1}^{n}\left[\log\left(\frac{k^2c^2 Nn\Delta x}{4\pi\nu_r^2}\right) - \log\left(\frac{1-\beta^2}{\beta^2}\right) - \beta^2\right]$$

$\ldots\ldots(6).$

Experiments on the absorption of homogeneous β-rays have been carried out by Varder (16). His curves for aluminium agree over considerable parts of their length with the linear law previously found by Wilson. By producing these linear portions to cut the x-axis, Varder estimated the ranges of β-rays of various velocities in aluminium. His results agree with equation (5) over a wide range of velocities to within about 20 per cent. The ranges of cathode rays in a number of metals have been measured by Schonland (17), whose results for aluminium show a substantial agreement with Bohr's theory. The reduction in velocity of cathode rays has been measured by Terrill (18), who verified the Thomson-Whiddington law and found the constant entering to be of the order of magnitude predicted by Bohr.

The experiments on rays of velocities from $0{\cdot}2C$ to $0{\cdot}99C$ have yielded results which substantiate Bohr's theory as to orders of magnitude, but which differ to a considerable extent among themselves. In the present experiments the writer has endeavoured to eliminate some sources of error which enter into the measurement of R and $\Delta(H\rho)$ for β-rays.

Digitized by Google

Measurements of the ranges of β-rays in aluminium. The apparatus is shown in Fig. 1. The box containing the source A of radiation is placed between the pole pieces of an electromagnet so that those particles which are projected in a direction normal to the magnetic field describe circular paths of curvature depending on their velocities. A limited beam of mean radius of curvature $\rho = 6$ cms. passes through the slot B and is brought to a focus on to a mica window C of 1·8 cms. stopping power for α-particles and width 2 mms., where it leaves the box and enters the ionisation chamber D. The γ-radiation is absorbed as far as possible by the lead blocks H and J, and its effect in the ionisation chamber is still further reduced by means of the balance chamber E. The walls of the ionisation chamber are charged positively, those of the balance chamber negatively, and the opposing ionisation currents

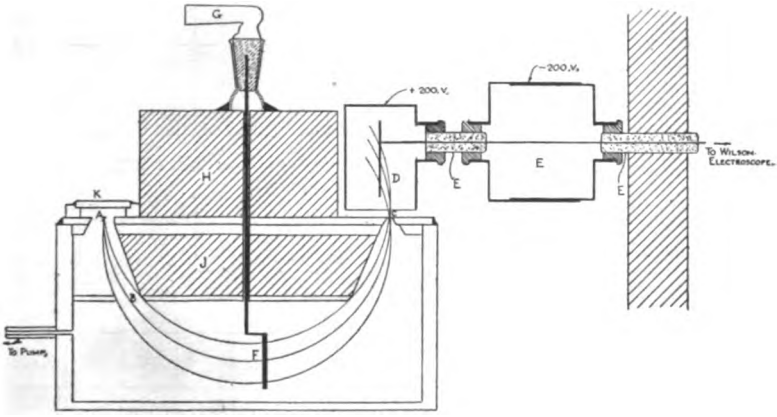


Fig. 1.

due to γ-radiation are nearly balanced out by adjusting the size of E. The ionisation current is measured by means of a Wilson tilted electroscope working at sensitivities of from 30 to 200 divisions per volt. The source, consisting of a tube containing radium emanation, is placed in position by removing the plate K, and the box is evacuated by means of a Gaede pump.

The ionisation in D may be due to three causes, viz. (1) the direct β-radiation, (2) the scattered β-radiation, and (3) the unbalanced γ-radiation. To separate out the three effects a device previously used by Chadwick (19) was employed. This consists of a lead screen F which can be interposed in or removed from the path of the direct beam by turning G. When the direct beam is free to enter the ionisation chamber all three effects are measured together. If now the screen F be interposed in the path of the direct beam the ionisation current is the result of (2) + (3);

Digitized by Google

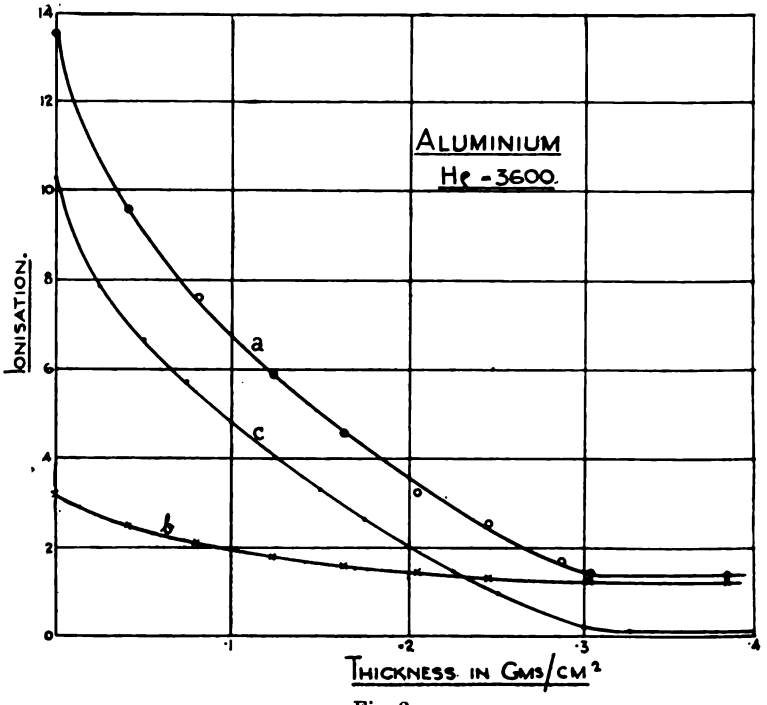


Fig. 2.

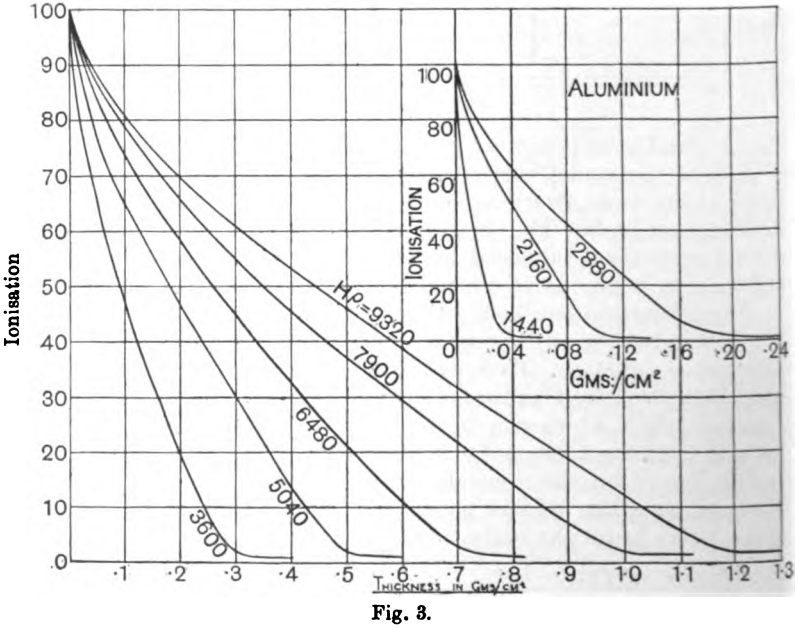


Fig. 3.

Digitized by Google

and when the exciting current of the magnet is reversed, of (3) alone. The effect of the direct beam is thus represented by the difference between the readings corresponding to the two positions of *F*. In Varder's experiments no attempt was made to allow for the scattered β-radiation, though steps were taken to diminish it. To what extent this would affect his results is uncertain, but it is found in these experiments that in measurements on the high-

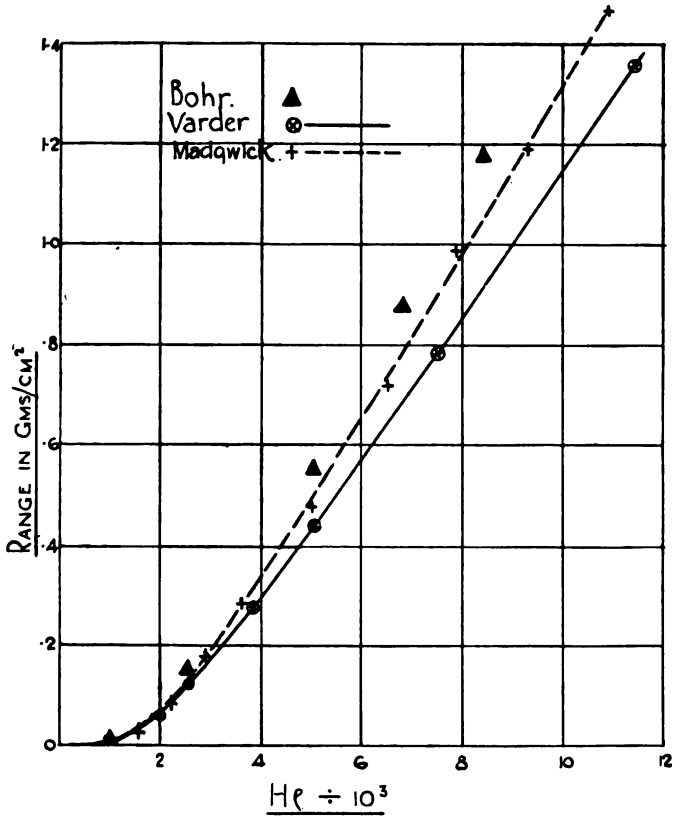


Fig. 4.

velocity rays the correction for scattered radiation is very appreciable. In an investigation of the continuous β-ray spectrum of Ra ($B + C$) the writer found the highest velocity rays to correspond to about 12,000 $H\rho$. Effects which Varder detected up to 16,000 $H\rho$ were probably due solely to scattered radiation.

The magnetic field was calibrated against the current through the electromagnet by comparison with Ellis's values (20) for the

Digitized by Google

principal lines in the spectrum of Ra *B*. The field being adjusted to a particular value, sheets of absorbing matter of progressively increasing thickness were placed over the slit *C* and measurements of the ionisation currents were made in each case. The results of a typical experiment on aluminium are shown in Fig. 2. The curve *a* represents the effects of the direct β-radiation, the scattered β-radiation, and the unbalanced γ-radiation, *b* that of the scattered

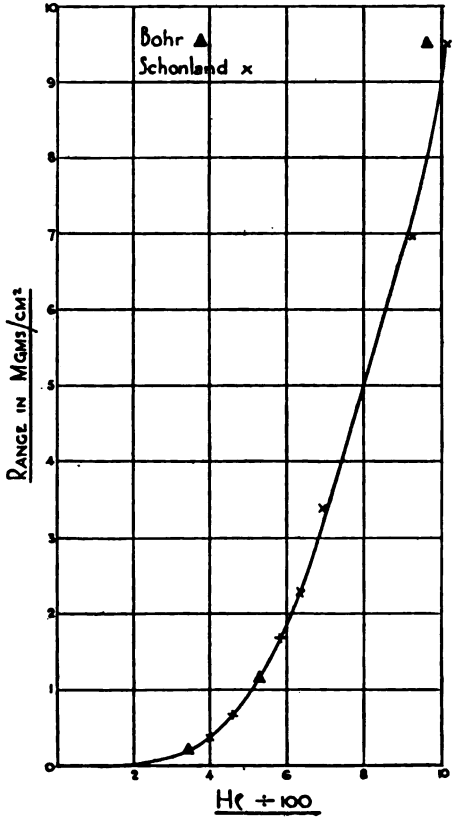


Fig. 5.

β plus the unbalanced γ-radiation, and *c* the difference between the two. Fig. 3 shows a series of curves *c* corresponding to various values of $H\rho$. It will be seen that they agree with the results of Wilson and Varder in that they are very nearly linear over a considerable part of their length. Values of the ranges as estimated by producing these linear portions to cut the x-axis are set out in Fig. 4 together with the results of Varder's experiments and some

Digitized by Google

theoretical values calculated from equation (5). It is of interest to note that the values of the ranges which are given by the constant a of equation (3) in Wilson's early experiments show a good general agreement with the theory. For the sake of completeness Schonland's results are given in Fig. 5.

Absorption experiments were also carried out on copper, tin, and lead. These gave a progressive deviation from the linear law of absorption due to the increasing scattering effect.

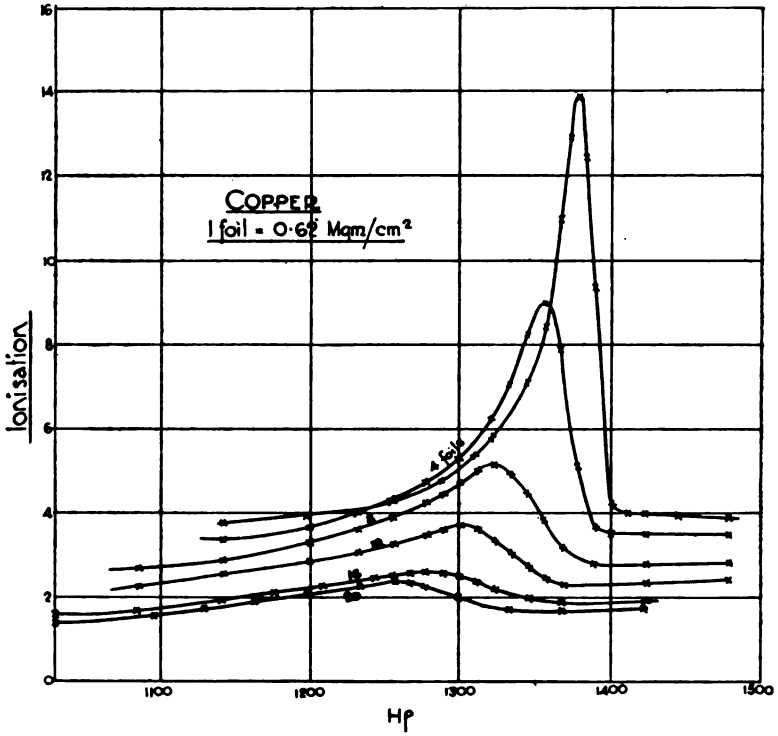


Fig. 6.

Experiments on the reduction in velocity of homogeneous β-rays. No great accuracy can be claimed for measurements on the reduction in velocity of β-rays by the photographic method. The lines are displaced only slightly and are so broadened that it is difficult to estimate the position of maximum intensity. On the other hand, experiments can be carried out on a number of lines at the same time.

The distribution of intensity in the displaced and broadened line can be investigated accurately by the electrical method, provided

Digitized by Google

that the original line is sufficiently pronounced. The Th B line of $H\rho$ 1382 is one of exceptional intensity and this has been examined before and after its passage through various thicknesses of aluminium, copper, silver, and gold. The source was of Th $(B + C)$ on a platinum plate, the back of which was enamelled. The plate was about 2 mms. wide and was inclined to the vertical (i.e. to the initial direction of the beam) at an angle of 29°. The effective thickness

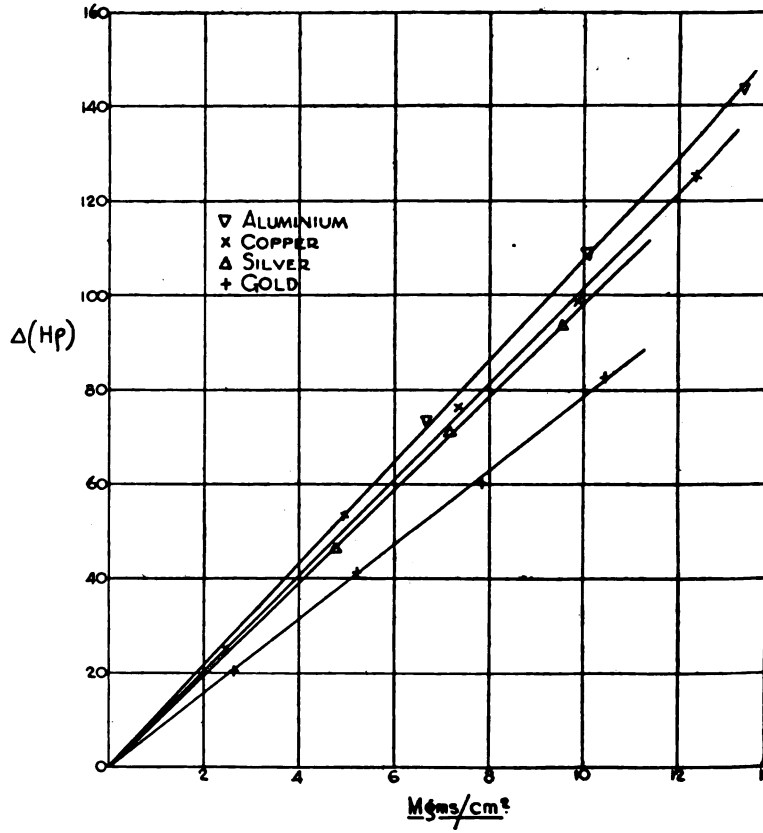


Fig. 7.

of the absorbing sheet is thus cosec 29° = 2·06 times its actual thickness. The sources were of the order of one or two mgms. as measured by their γ-ray activity, and for such small sources the balance chamber was unnecessary. The results of a typical experiment for copper are shown in Fig. 6 and the displacements of the maxima of the peaks in Fig. 7. The value of $\Delta(H\rho)$ for a given mass per sq. cm. is distinctly less for gold than for aluminium,

Digitized by Google

which is at variance with the conclusion reached by Danysz, Terrill, d'Espine and others that the reduction in velocity per unit

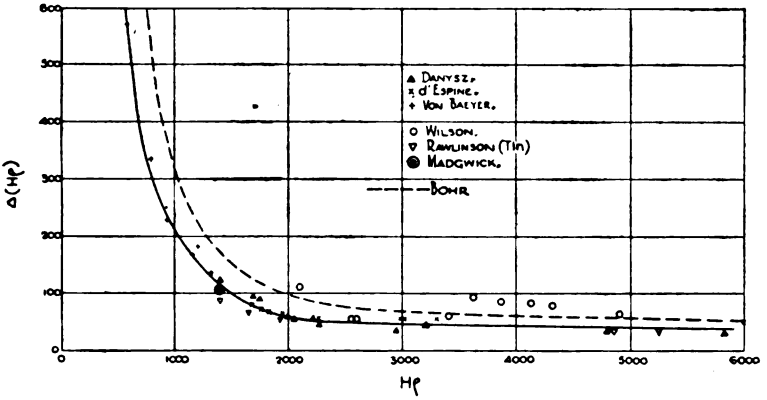


Fig. 8.

mass per sq. cm. is constant. For aluminium the value of $\Delta(H\rho)$ for a thickness of 0·01 gm. per sq. cm. was found to be 107 whereas

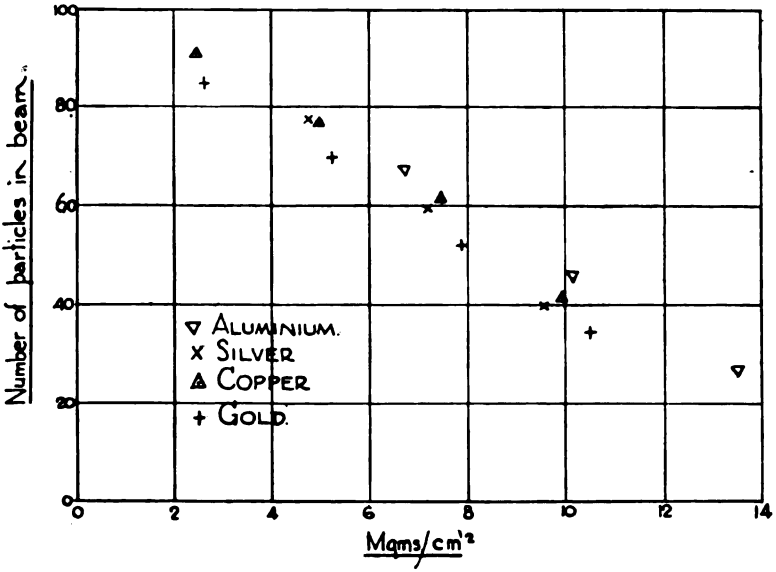


Fig. 9.

the value calculated from equation (6) is 171. The values of this quantity obtained from the results of other experimenters, together with Bohr's theoretical curve, are shown in Fig. 8.

Digitized by Google

A rough estimate of the relative numbers of particles remaining in the beam has been formed by detaching the peaks from the continuous background and determining their areas after correcting for change in $H\rho$, the variation in ionisation due to change in velocity, and the decay of the source. The results are shown in Fig. 9.

Conclusion. There is a close parallelism over a very wide range of energies between the theoretical and experimental results both for R and $\Delta(H\rho)$, which lends strong support to the general truth of the theory. Considering the complexity of β-ray absorption and hence the uncertainty attaching to the interpretation of the experimental data, the disparity between theory and experiment is not great. The ranges measured by Varder and the writer are less than those deduced from equation (5), but no account is taken by Bohr of the tortuosity of the path described by the β-particle. C. T. R. Wilson (21), by his cloud method, has measured the total ranges of electrons ejected from atoms by X-rays, and has found values which are greater than those calculated from the theory. The experimental values of $\Delta(H\rho)$ are nearly all decidedly less than the theoretical values, but here again it is to be noted that increasing numbers of particles are withdrawn from the beam and the experimental results do not represent the most probable loss of energy in the whole beam.

It is a pleasure to thank Professor Sir Ernest Rutherford for his kind interest in these experiments. I am also indebted to Dr J. Chadwick for the sources of Th$(B+C)$ and for much helpful advice; to Mr G. R. Crowe for preparing the emanation tubes; and to the Department of Scientific and Industrial Research for financial assistance.

REFERENCES.

(1) Lenard, *Wied. Ann.* 56, p. 255 (1895).
(2) Leithauser, *Ann. d. Phys.* 18, p. 410 (1905).
(3) Becker, *Ann. d. Phys.* 17, p. 445 (1905).
(4) Rutherford, *Phil. Mag.* 47, p. 109 (1899).
(5) Schmidt, *Phys. Zeit.* p. 764 (1906); p. 361 and p. 737 (1907); p. 929 (1909).
(6) Hahn and Meitner, *Phys. Zeit.* p. 321 (1908); p. 948 (1909).
(7) McClelland, *Trans. Roy. Soc. Dubl.* (1905, 1907).
(8) Schmidt, *Ann. d. Phys.* 23, p. 671 (1907).
(9) Whiddington, *Proc. Roy. Soc.* A, 86, p. 360 (1911–12); 89, p. 554 (1913–14).
(10) Wilson (W.), *Proc. Roy. Soc.* A, 82, p. 612 (1909); 84, p. 141 (1910).
(11) Danysz, *Le Radium*, p. 1 (1912).

Digitized by Google

(12) VON BAEYER, *Phys. Zeit.* p. 485 (1912).
(13) RAWLINSON, *Phil. Mag.* 30, p. 627 (1915).
(14) D'ESPINE, *Comptes Rendus*, p. 458 (1926).
(15) BOHR, *Phil. Mag.* 25, p. 10 (1913); 30, p. 581 (1915).
(16) VARDER, *Phil. Mag.* 29, p. 725 (1915).
(17) SCHONLAND, *Proc. Roy. Soc.* A, 104, p. 235 (1923); 108, p. 187 (1925).
(18) TERRILL, *Phys. Rev.* 22, p. 101 (1923).
(19) CHADWICK, *Verh. d. D. Phys. Ges.* 16, p. 383 (1914).
(20) ELLIS, *Proc. Roy. Soc.* A, 101, p. 1 (1922).
(21) WILSON (C. T. R.), *Proc. Roy. Soc.* A, 104, p. 1 and p. 192 (1923).

Digitized by Google

The β-ray spectrum of Ra E. By E. MADGWICK, M.C., Ph.D., Emmanuel College. (Communicated by Prof. Sir ERNEST RUTHERFORD.)

[*Received* 22 August, *read* 24 October 1927.]

Introduction.

In the course of an investigation of the β-ray spectrum of Ra ($B + C$) by the ionisation method, Chadwick (1) concluded that the line spectrum is superimposed on a strong continuous background. The ionisation method, however, is not one which admits of high resolving power, and doubts have been entertained whether the continuous spectrum has in fact any real existence. In this connection the case of Ra *E* is one of special importance, since no line spectrum from this body has been detected by the more sensitive photographic method. Experiments have therefore been undertaken with the object of determining the distribution with velocity of the numbers of particles in the β-ray spectrum of Ra *E*.

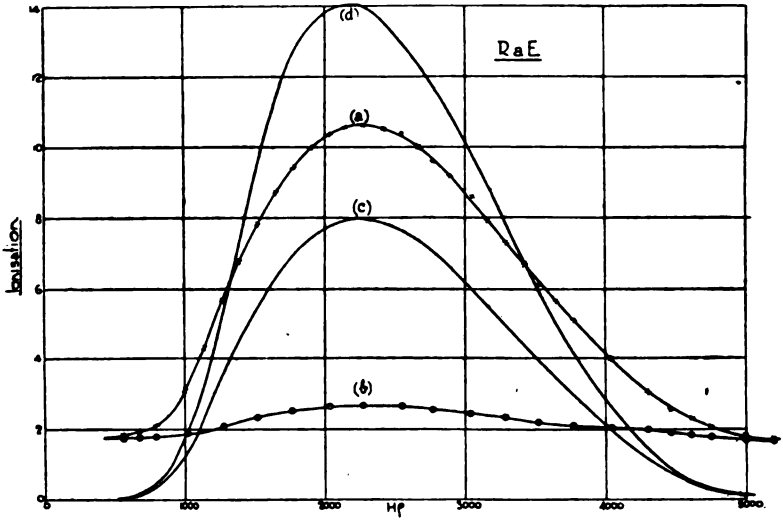
Experimental. The source of radiation is contained in a brass box which is situated between the pole pieces of an electromagnet and evacuated to a pressure of a fraction of a millimetre. A limited beam of the rays of radius of curvature $\rho = 6$ cms. is focussed on to a mica window where it passes out of the box and enters an ionisation chamber, the ionisation current being measured by means of a Wilson tilted electroscope working at a sensitivity of 200 divisions per volt. By adjusting the current through the electromagnet the ionisation currents corresponding to successive values of $H\rho$ are determined. To eliminate the effect of scattered radiation a small lead screen, sufficiently thick to stop the β-rays, is arranged so that it can be turned into or out of the path of the direct beam so that the effects of the scattered and of the direct + scattered radiation can be determined separately, the difference representing the effect of the direct rays. The source was obtained by rotating a small nickel plate in a solution of Ra *D*. The plate was 5 mms. long and 3 mms. wide and was arranged so that the beam to be measured left the plate at a very small angle.

Experimental results are shown in the diagram. The curve (*a*) represents the effect of the direct + scattered radiation, (*b*) that of the scattered radiation alone, and (*c*), the difference between the two, that of the direct beam. To estimate the distribution of the numbers of particles with $H\rho$ the curve (*c*) must be corrected in two respects. The beam contains particles of radii of curvature between ρ and $\rho + \delta\rho$. Therefore for any field H the rays have values of $H\rho$ between $H\rho$ and $H(\rho + \delta\rho)$, that is a range of $H\delta\rho$. In order that each reading shall correspond to a

Digitized by Google

constant increment of $H\rho$ it is thus necessary to divide it by H. Also, the numbers of pairs of ions produced by an electron per cm. of path varies with the velocity, and for fairly low velocities the variation is very considerable. Block (2) has collected data from all available sources and gives a curve showing the number of ions produced per centimetre path in air by particles of velocities up to $2{\cdot}9 \times 10^{10}$ cms. per sec. This is employed to estimate from the ionisation currents the relative numbers of particles entering the ionisation chamber. Corrected in these respects the results are represented by the curve (*d*).

Discussion. The curve starts at about 600 $H\rho$, rises to a maximum at about 2200 $H\rho$, and ends at about 5000 $H\rho$. The



beginning will be affected to some extent by absorption of the rays in the mica window (of 1·8 cms. stopping power for α-particles) and the 2 mms. of air space between the window and the ionisation chamber. It is to be noted also that the rays which are measured include those reflected from the nickel plate, which would tend to move the maximum in the direction of lower velocity. There is no reason to doubt the accuracy of the end point.

Ellis (3) considers that the particles comprising the continuous spectrum are ejected from the nucleus through a wide range of velocities. Meitner (4), on the other hand, supposes them to be ejected with constant velocity. If Meitner's supposition is correct it is clear that varying amounts of energy are lost by the particles

Digitized by Google

during their passage out of the atoms; the energies represented by the end point and by the maximum of the spectrum differ by roughly 7×10^5 volts. If this loss is to be attributed to collisions with extranuclear electrons the number of particles ejected must be substantially greater than one per disintegrating atom, which is in conflict with the results of Emeléus's experiments (5); and there ought to be present κ-radiation, which, however, is not found in any quantity. Rosseland (6) suggests that the continuous spectrum may be due to the radiation of energy as the β-ray passes through the intense fields within the atom, but Ra *E* emits practically no γ-radiation. The phenomena attaching to the disintegration of Ra *E* are relatively simple, yet by virtue of their simplicity they impose restrictions on the way of explaining the continuous β-ray spectrum; and at present there appears to be no alternative but to follow Ellis and postulate that the particles are ejected from the nucleus with velocities which vary continuously.

The continuous spectra of Ra *C* and Th $(B+C)$ have also been investigated. The values for the maxima and end points are set out below.

	Maxima		End points	
	$H\rho$	volts	$H\rho$	volts
Ra *C*	2500	4×10^5	12,000	30×10^5
Ra *E*	2200	$3{\cdot}26 \times 10^5$	5,000	$10{\cdot}7 \times 10^5$
Th $(B+C)$	1400	$1{\cdot}52 \times 10^5$	8,800	22×10^5

I am greatly indebted to Prof. Sir Ernest Rutherford for his interest in the experiment; to Dr J. Chadwick for providing the source; and to the Council of Armstrong College for financial assistance.

REFERENCES.

(1) Chadwick, *Verh. d. D. Phys. Ges.* 16, p. 383 (1914).
(2) Block, *Ann. d. Phys.* 38, p. 559 (1912).
(3) Ellis, *Proc. Roy. Soc.* A, 101, p. 1 (1922).
(4) Meitner, *Z. f. Phys.* (1922, 1923, 1924).
(5) Emeléus, *Proc. Camb. Phil. Soc.* 22, p. 400 (1924).
(6) Rosseland, *Z. f. Phys.* 14, p. 173 (1923).

Digitized by Google

PROCEEDINGS AT THE MEETINGS HELD DURING THE SESSION 1926—1927.

ANNUAL GENERAL MEETING.

October 25, 1926.

In the Cavendish Laboratory.

PROFESSOR J. T. WILSON, PRESIDENT, IN THE CHAIR.

The following were elected Officers for the ensuing year:

President:

Dr Horace Lamb.

Vice-Presidents:

Mr G. Udny Yule.
Prof. J. T. Wilson.
Prof. A. Hutchinson.

Treasurer:

Mr F. A. Potts.

Secretaries:

Mr F. P. White.
Mr R. H. Fowler.
Mr H. Munro Fox.

Members of the Council who continue:

Dr C. D. Ellis.
Mr F. F. Blackman.
Prof. C. T. R. Wilson.
Mr J. E. Littlewood.
Prof. Nuttall.
Dr J. Chadwick.
Mr W. B. R. King.
Mr P. M. S. Blackett.

New Members of Council:

Prof. T. M. Lowry.
Mr H. Jeffreys.
Dr F. J. W. Roughton.
Mr F. T. Brooks.

The President announced the deaths of the following Fellows, since the annual General Meeting 1925:

Dr William Bateson, St John's College.
Mr E. J. Bles, King's College.
Mr C. G. F. James, Trinity College.
Prof. J. N. Langley, Trinity College.
Prof. W. J. Lewis, Trinity College.

Digitized by Google

The following were elected Associates of the Society:

R. S. Bartlett.
F. W. Bradley, Trinity College.
Max Brunner.
E. C. Bryant.
L. W. Elder, Fitzwilliam House.
B. A. Fletcher, Sidney Sussex College.
R. P. Gillespie, St John's College.
T. Iwatsuki.
G. Jessop, Trinity College.
O. A. Jones, Emmanuel College.
C. T. Knipp.
W. V. Lloyd, Emmanuel College.
C. A. MacConkey, St John's College.
Miss S. M. Manton, Girton College.
E. F. Porter, Trinity Hall.
Mrs E. Salaman, Newnham College.
Miss K. I. Sayers, Girton College.
J. H. Schulman, Trinity Hall.
G. F. Smith, Fitzwilliam House.
K. Suda.
W. Taylor, Trinity College.
E. E. Watson, Gonville and Caius College.

The following Communications were made to the Society:

1. On compressional waves in two superposed layers. By Mr H. JEFFREYS, St John's College.

2. (1) On the descriptive form of Taylor's theorem.
(2) The summation of a Fourier integral of finite type.

By Mr S. POLLARD, Trinity College.

3. Note on the Petzval optical condition. By Mr G. C. STEWARD, Gonville and Caius College.

4. Maxwell's law, and the absorption and emission of radiation. By Mr E. A. MILNE, Trinity College.

5. The correspondence between lines in threefold space and points of a quadric fourfold in fivefold space, established by a geometrical construction. By Mr T. L. WREN, St John's College.

6. Improvements in the apparatus for measuring the velocity of very rapid chemical reactions. II. By Dr ROUGHTON, Trinity College, and Dr HARTRIDGE, King's College.

7. On the stability of copper catalysts produced by thermal decomposition. By Dr CONSTABLE, St John's College.

8. Note on the velocity of gas-reactions. By J. A. CHRISTIANSEN. (Communicated by Mr R. H. FOWLER.)

9. Note on the effect of alpha particles on paraffin. By W. T. RICHARDS. (Communicated by Professor Sir E. RUTHERFORD.)

10. On the figure of Pappus' theorem. By C. V. HANUMANTA RAO. (Communicated by Professor H. F. BAKER.)

Digitized by Google

November 8, 1926.

In the Botany School.

DR HORACE LAMB, PRESIDENT, IN THE CHAIR.

The following were elected Fellows of the Society:

J. A. Ratcliffe, B.A., Sidney Sussex College.
F. P. Smith, B.A., Peterhouse.
A. F. H. Ward, B.A., Jesus College.

The following were elected Associates of the Society:

T. E. Allibone, Gonville and Caius College.
O. M. B. Bulman, Sidney Sussex College.
H. M. Cave, Trinity College.
R. R. Nimmo, Clare College.
J. V. L. Rennie, Emmanuel College.

The following Communications were made to the Society:

1. Recent advances in Stereochemistry. By Professor T. M. LOWRY.

2. Some investigations of gas discharges by means of an exploring electrode. By Dr K. G. EMELÉUS, St John's College. (Communicated by Mr E. V. APPLETON.)

3. The absorption spectra of "saturated" and "unsaturated" organic substances. By Mr J. E. PURVIS, Corpus Christi College.

November 22, 1926.

In the Cavendish Laboratory.

DR HORACE LAMB, PRESIDENT, IN THE CHAIR.

The following were elected Fellows of the Society:

J. G. Adshead, B.A., Gonville and Caius College
E. T. S. Appleyard, B.A., King's College.
A. R. Clapham, B.A., Downing College.
P. I. Dee, B.A., Sidney Sussex College.
P. A. M. Dirac, Ph.D., St John's College.
E. P. Hudson, B.A., King's College.
M. H. Salaman, B.A., Trinity College.
L. H. Thomas, B.A., Trinity College.
M. A. Vernon, B.A., Trinity College.
O. H. Wansbrough-Jones, B.A., Trinity Hall.
A. H. Wilson, B.A., Emmanuel College.

The following Communications were made to the Society:

1. On the polarisation of mercury lines emitted from a discharge tube in a magnetic field. By Dr H. W. B. SKINNER, Trinity College.

2. The application of the method of the magnetic spectrum to the study of secondary electronic emission. By C. F. SHARMAN, B.A., King's College. (Communicated by Professor Sir E. RUTHERFORD.)

Digitized by Google

3. The calculation of atomic fields. By L. H. THOMAS, B.A., Trinity College.

4. A contribution to the theory of diffusion pumps. By L. WERTENSTEIN. (Communicated by Professor Sir E. RUTHERFORD.)

5. On a group of order 25920 and the projective transformations of a cubic surface. By Mr W. BURNSIDE, Pembroke College.

6. The Compton effect in wave mechanics. By Dr P. A. M. DIRAC, St John's College.

7. A mathematical theory of natural and artificial selection. Part IV. By Mr J. B. S. HALDANE, Trinity College.

8. On the addition of the primary aberrations. By Mr G. C. STEWARD, Gonville and Caius College.

January 17, 1927.

In the School of Agriculture.

DR HORACE LAMB, PRESIDENT, IN THE CHAIR.

The following were elected Fellows of the Society:

S. J. Hickson, M.A., Downing College.
J. Hilton, B.A., Christ's College.
W. H. McCrea, B.A., Trinity College.
E. B. Moullin, M.A., Downing College.

The following were elected Associates of the Society:

F. B. Bowden, Gonville and Caius College.
Miss L. Chitty, Newnham College.
B. H. C. Turvey, St Catharine's College.

The following Communication was made to the Society:

Animal Calorimetry. By Professor T. B. WOOD.

January 31, 1927.

In the Cavendish Laboratory.

DR HORACE LAMB, PRESIDENT, IN THE CHAIR.

The following Communications were made to the Society:

1. What determines the resistance and tilt of an aeroplane? By Professor Sir J. LARMOR.

2. The effect of the orbital velocity of the electrons in heavy atoms on their stopping of α particles. By L. H. THOMAS, B.A., Trinity College.

Digitized by Google

3. Surface adsorption and the velocity of chemical action at gas solid interfaces. By Dr F. H. CONSTABLE, St John's College.

4. The evaluation of Gibbs' phase integral for imperfect gases. By H. D. URSELL, Trinity College. (Communicated by Mr R. H. FOWLER.)

5. Note on Ramanujan's arithmetical function $\tau(n)$. By Mr G. H. HARDY, Trinity College.

6. Notes on the theory of series (iii): On the summability of the Fourier series of a nearly continuous function. By Mr G. H. HARDY, Trinity College, and Mr J. E. LITTLEWOOD, Trinity College.

7. The pedal (3, 2) correspondence. By R. VAIDYANATHASWAMY. (Communicated by Mr H. W. TURNBULL.)

8. Lines and planes in a metrical space. By W. HUNTER, B.A., Trinity College. (Communicated by Mr S. POLLARD.)

February 14, 1927.

In the Botany School.

DR HORACE LAMB, PRESIDENT, IN THE CHAIR.

The following were elected Fellows of the Society:

H. Lob, M.A., King's College.
F. I. G. Rawlins, M.Sc., Trinity College.
T. W. Wormell, B.A., St John's College.

The following were elected Associates of the Society:

W. R. Andress, Trinity College.
S. E. A. Landale, Corpus Christi College.

The following Communication was made to the Society:

Some problems in the Evolution of the Echinoidea. By Professor H. L. HAWKINS.

February 28, 1927.

In the Botany School.

DR HORACE LAMB, PRESIDENT, IN THE CHAIR.

The following Communications were made to the Society:

1. Some problems of plant biology in Malaya. By Mr R. E. HOLTTUM, St John's College.

2. The mammalian sex-ratio. By M. A. S. PARKES, Christ's College. (Communicated by Mr H. MUNRO FOX.)

3. On the present position of the mitochondria and the Golgi apparatus. By Dr V. NATH, Trinity Hall. (Communicated by Mr H. MUNRO FOX.)

Digitized by Google

4. Glutathione. By Mr H. E. TUNNICLIFFE, Gonville and Caius College. (Communicated by Mr H. MUNRO FOX.)

5. On Herschel's condition and the optical Cosine law. By Mr G. C. STEWARD, Gonville and Caius College.

March 14, 1927.

In the Cavendish Laboratory.

DR HORACE LAMB, PRESIDENT, IN THE CHAIR.

The following was elected a Fellow of the Society:

F. T. K. Pentelow, B.A., Christ's College.

The following Communications were made to the Society:

1. An experiment on the stability of superposed streams of fluid. By Mr G. I. TAYLOR, Trinity College.

2. The absolute intensities of the γ-rays of radium B and radium C. By Dr C. D. ELLIS, Trinity College, and W. A. WOOSTER, B.A., Peterhouse.

3. The limits of classical scattering. By Mr P. M. S. BLACKETT, King's College.

4. The stopping power of hydrogen atoms for α-particles according to the new mechanics. By J. A. GAUNT, B.A., Trinity College. (Communicated by Mr R. H. FOWLER.)

5. Absorption coefficients for X-rays. By L. H. MARTIN, Trinity College. (Communicated by Professor Sir E. RUTHERFORD.)

6. (1) On the expression of a rational integral function of a single variable in terms of two other such functions, which are prime to one another.

(2) On the fifth book of Euclid's elements (Second addendum to fifth paper).

By Dr M. J. M. HILL, Peterhouse.

May 2, 1927.

In the Cavendish Laboratory.

DR HORACE LAMB, PRESIDENT, IN THE CHAIR.

The following Communications were made to the Society:

1. Excitation by high velocity electrons. By Mr E. G. DYMOND, St John's College.

2. Synthesis of ammonia by electrons. By A. CARESS, B.A., Trinity Hall. (Communicated by Mr E. K. RIDEAL.)

3. The light intensity of the calcium chromosphere. By P. A. TAYLOR, B.A., Emmanuel College.

Digitized by Google

4. On the effect of the addition of successive small quantities of poisonous substances on the velocity of catalytic gas reactions in closed vessels. By Dr F. H. CONSTABLE, St John's College.

5. On Mathieu Functions. By S. GOLDSTEIN, B.A., St John's College.

6. A case of distinction between Fourier integrals and Fourier series. By Miss M. E. GRIMSHAW, B.A., Newnham College. (Communicated by Mr S. POLLARD.)

7. Wave propagation in strings with continuous and concentrated loads. By Mr H. JEFFREYS, St John's College.

May 16, 1927.

In the Botany School.

PROFESSOR J. T. WILSON, VICE-PRESIDENT, IN THE CHAIR.

The following were elected Fellows of the Society:

J. A. Chalmers, B.A., Queens' College.
W. L. Edge, B.A., Trinity College.

The following was elected an Associate of the Society:

S. B. Setna, Fitzwilliam House.

The following Communications were made to the Society:

1. The carbohydrate metabolism of the developing frog embryo. By Dr J. NEEDHAM, Gonville and Caius College.

2. Some structural characters of the genus *Dictyonema* Hall, and the technique employed in their determination. By O. M. B. BULMAN, Sidney Sussex College. (Communicated by Mr W. B. R. KING.)

3. Some problems in the comparison of chromosomes. By J. S. YEATES, Trinity College. (Communicated by Mr F. T. BROOKS.)

4. Les théories de la polarité dans les phénomènes de régénération. By MARCEL ABELOOS. (Communicated by Mr H. MUNRO FOX.)

5. The conditions governing parturition. By Dr F. H. A. MARSHALL, Christ's College.

6. The mechanics of vertebrate development. By G. R. DE BEER. (Communicated by Mr H. MUNRO FOX.)

July 25, 1927.

In the Cavendish Laboratory.

DR HORACE LAMB, PRESIDENT, IN THE CHAIR.

The following were elected Fellows of the Society:

A. Caress, B.A., Trinity Hall.
D. Stockdale, Ph.D., King's College.

Digitized by Google

The following were elected Associates of the Society:

D. Barnett, Corpus Christi College.
Miss S. F. Cox.
G. H. Dickson, Corpus Christi College.
Miss N. Martin, Girton College.
R. J. Pumphrey, Trinity Hall.
Miss D. H. Strangeways, B.A.

The following Communications were made to the Society:

1. A valve amplifier for ionisation currents. By C. E. WYNN-WILLIAMS, Trinity College. (Communicated by Professor Sir E. RUTHERFORD.)

2. A differential retarding potential method for the study of the energy distribution of slow electron emissions. By C. F. SHARMAN, B.A., King's College. (Communicated by Professor Sir E. RUTHERFORD.)

3. The Hall effect in single crystals of iron. By Dr W. L. WEBSTER, Trinity College. (Communicated by Dr KAPITZA.)

4. The production of characteristic X-rays by electronic impact. By Dr L. H. THOMAS, Trinity College.

5. (1) A mathematical theory of natural and artificial selection. Part V: selection and mutation.

(2) The comparative genetics of colour in Rodents and Carnivora.

By Mr J. B. S. HALDANE, Trinity College.

6. On the generalizations of the theorems of Parseval and Riesz-Fischer. By Mr S. POLLARD, Trinity College.

7. Eddy currents in thin circular cylinders of uniform conductivity due to periodically changing magnetic fields, in two dimensions. By Dr F. W. CARTER, St John's College.

8. (1) A theorem of Dr P. Zeeman.

(2) The pedal planes of a tetrahedron.

(3) An illustration of the space representation of circles.

By Mr J. H. GRACE, Peterhouse.

9. Summation of the integral conjugate to the Fourier integral of finite type. By Miss M. E. GRIMSHAW, Newnham College. (Communicated by Mr S. POLLARD.)

10. Simplexes and other configurations upon a rational normal curve. By Mr F. P. WHITE, St John's College.

11. Notes on spectrophotometry. By R. W. DITCHBURN, B.A., Trinity College.

12. The vertical distribution of plankton in the sea. By F. S. RUSSELL. (Communicated by Mr H. MUNRO FOX.)

13. On some fundamental laws of ovarian dynamics. By A. LIPSCHÜTZ. (Communicated by Mr H. MUNRO FOX.)

Digitized by Google

STATEMENT OF ACCOUNTS

Digitized by Google

STATEMENT OF ACCOUNTS OF THE CAMBRIDGE PHILOSOPHICAL SOCIETY, FROM JANUARY 1, 1926, TO DECEMBER 31, 1926.

Receipts.

	£ s. d.	£ s. d.
Balance due to the Society, Jan. 1, 1926		
On petty cash account ...		1 7 11
Subscriptions for 1926		484 1 0
Associates' Subscriptions for 1926 ...		79 5 6
Tickets for Library		14 15 0
Sale of Publications		8 1 8
University Press		196 6 7
Dividends on Liverpool Corporation Stock	8 12 0	
Dividends on Conversion Loan ...	26 16 2	
		35 8 2
Interest on Deposit Account 1926 ...		17 6 6
Transferred from Dinner Fund ...		5 9 10
		£842 2 2

Expenditure.

	£ s. d.	£ s. d.
Balance due from the Society, Dec. 31, 1925		48 14 4
University Press		28 6 0
„ „ (Clay)		717 9 5
National Research Council		7 4 0
Grant for Lantern Slides (Fox) ...		1 3 0
Carriage of parcels and postage ...		10 18 1
Sundries		1 0
Balances in hand, Dec. 31, 1926:		
On current account	26 16 6	
With Treasurer: On petty cash account	1 9 10	
		28 6 4
		£842 2 2

SUSPENSE ACCOUNT.

Receipts.

	£ s. d.	£ s. d.
Balance due to the Account, Jan. 1, 1926		56 12 0
Composition Fee		21 0 0
Sale of Publications:		
Brown University	62 14 0	
Washington University	84 0 0	
		146 14 0
Sale of Books (Hoffer and Deighton) ...		11 5 0
		£235 11 0

Expenditure.

	£ s. d.
Balance due to the Account, Dec. 31, 1926	235 11 0
	£235 11 0

Digitized by Google

Receipts.	£	s.	d.
Balance due to the Fund, Jan. 1, 1926	352	18	8
Dividends on £750 L.N.E.R. 4 per cent. Debentures:			
January, 1926	12	0	0
July, 1926	12	0	0
	£376	18	8

Expenditure.	£	s.	d.
Balance due to the Fund, Dec. 31, 1926	376	18	8
	£376	18	8

STATEMENT OF BALANCES.

Receipts.	£	s.	d.	£	s.	d.
Balance at Messrs Barclays, Dec. 31, 1926: On Deposit	756	18	8			
Less due from the Society on current account overdrawn	117	12	6			
				£639	6	2

Expenditure.	£	s.	d.
Balance due to Hopkins Prize Fund Account	376	18	8
Balance due to Society on Suspense Account	235	11	0
Balance due to Society on Current Account	26	16	6
	£639	6	2

CAPITAL ACCOUNT (RESERVE FUND).

£307. 1*s.* 10*d.* Liverpool Corporation 3½ per cent. Stock.
£957. 7*s.* 8*d.* Conversion Loan 3½ per cent. @ 76¼.

F. A. POTTS, *Treasurer.*

REPORT OF THE AUDITORS.

We have examined the Treasurer's accounts and find them correct. We have also verified the investments of £307. 1*s.* 10*d.* Liverpool Corporation Stock, £957. 7*s.* 8*d.* Conversion Loan, £750 L.N.E.R. Co. 4 per cent. Debenture Stock.

28 *April* 1927.

E. C. FRANCIS
HAMILTON McCOMBIE } *Auditors.*

Digitized by Google

INDEX TO THE PROCEEDINGS VOL. XXIII

with references to the Transactions and Biological Reviews

ABELOOS, M., Les théories de la polarité dans les phénomènes de régénération. *See Biol. Rev.* II

Aberrations, primary, addition of the (STEWARD), 584

Absorption-lines, breadths of (DITCHBURN), 78

Absorption spectra of "saturated" and "unsaturated" organic substances (PURVIS), 588

Absorption of X-rays of long wave-length (MARTIN), 783

Absorption and reduction in velocity of β-rays on their passage through matter (MADGWICK), 970

Accounts, statement of, for 1925, 490; 1926, 994

Adiabatic hypothesis for magnetic fields (DIRAC), 69

ADSHEAD, J. G., Elected Fellow 1926, November 22, 987

Adsorption, surface (CONSTABLE), 593

Aeroplane, what determines the resistance and the tilt of an (LARMOR), 617

Affine field, integral invariants of the (NEWMAN), 262

Algebraic curves, plane, higher singularities of (SCOTT), 206

Aliphatic poly-amines, metallic complexes with the (MANN and POPE), 183

Alkali metals, distillation of (CLARK), 953

ALLIBONE, T. E., Elected Associate 1926, November 8, 987

ALLOT, E. N., Elected Associate 1925, October 26, 482

Alpha particles, effect of the orbital velocity of the electrons in heavy atoms on their stopping of (THOMAS), 713

Alpha particles on paraffin, effect of (RICHARDS), 516

Alpha particles, stopping power of hydrogen atoms for, according to the new quantum theory (GAUNT), 732

ANDRESS, W. R., Elected Associate 1927, February 14, 989

Angle, obtuse, hypothesis of the (HILL), 2

APPLETON, E. V., On the diurnal variation of ultra-short wave wireless transmission, 155

APPLEYARD, E. T. S., Elected Fellow 1926, November 22, 987

APPLEYARD, E. T. S., and H. W. B. SKINNER, A case of double reflexion, 951

ASTON, G. H., The amount of energy emitted in the γ-ray form by radium E, 935

ASTON, R. L., The tensile deformation of large aluminium crystals at crystal boundaries, 549

—— Some developments in the X-ray analysis of single crystals, 561

Atom, total energy of binding of a heavy (MILNE), 794

Atomic cores, polarisabilities of (SWIRLES), 403

Atomic fields, calculation of (THOMAS), 542

Atoms, hydrogen, stopping power of, for α-particles according to the new quantum theory (GAUNT), 732

Atoms, optical spectra of different, of the same electronic structure, II (HARTREE), 304

Atoms, recoil, photographic method of determining the mobility of (BRIGGS), 73

Axial rotation, relativity in connexion with (HARGREAVES), 198

BAKER, H. F., The stability of rotating masses of liquid; note to a former paper, 1

—— The reciprocation of one quadric into another, 22

—— On chains of two-two relations and the theory of elliptic functions, 92

BARNETT, D., Elected Associate 1927, July 25, 992

BARNETT, M. A. F., and J. A. RATCLIFFE, On the attenuation of wireless signals in short distance overland transmission, 288

BARTLETT, R. S., Elected Associate 1926, October 25, 986

BATESON, G., Elected Associate 1925, October 26, 482

BATESON, W., *Deceased*, 985

BEER, G. R. DE, The mechanics of vertebrate development. *See Biol. Rev.* II

Beta ray spectrum of Ra E (MADGWICK), 982

Beta rays on their passage through matter, absorption and reduction in velocity of (MADGWICK), 970

Digitized by Google

BIDDER, A. M., Elected Associate 1925, October 26, 482
BLACKETT, P. M. S., The limits of classical scattering, 698
BLES, E. J., *Deceased*, 985
BOHR, N., On the application of the quantum theory to atomic structure. Part I. *See Supplement*
BOWDEN, F. B., Elected Associate 1927, January 17, 988
BRACELIN, P., The period of decay of radium B and radium C, 150
BRADLEY, F. W., Elected Associate 1926, October 25, 986
BRIGGS, G. H., A photographic method of determining the mobility of recoil atoms, 73
BRUNNER, MAX, Elected Associate 1926, October 25, 986
BRYANT, E. C., Elected Associate 1926, October 25, 986
BULMAN, O. M. B., Elected Associate 1926, November 8, 987
—— Some structural characters of the genus *Dictyonema* Hall, and the technique employed in their determination. *See Biol. Rev.* II
BURKILL, J. C, On Mellin's inversion formula, 356
BURNETT, D., Elected Fellow 1926, January 18, 485
—— The relation between refractive index and density, 907
BURNSIDE, W., On a group of 1440 birational transformations of four variables that arises in considering the projective equivalence of double sixes, 103
—— On a group of order 25920 and the projective transformations of a cubic surface, 498

CANNON, H. G., Elected Fellow 1926, July 26, 488
Carbon dioxide and the form of the CO_2 molecule, specific heat of (MCCREA), 890
CARESS, A., Elected Associate 1926, July 26, 488; Elected Fellow 1927, July 25, 991
CARTER, F. W., Eddy currents in thin circular cylinders of uniform conductivity, due to periodically changing magnetic fields, in two dimensions, 901
CASTLE, E. S., Elected Associate 1925, October 26, 482
Catalysts, copper, produced by thermal decomposition, stability of (CONSTABLE), 432
Catalytic reaction, velocity of gaseous (CONSTABLE), 832
CATON, C. T. L., Elected Fellow 1926, February 1, 486
CAVE, H. M., Elected Associate 1926, November 8, 987
CHALMERS, J. A., Elected Fellow 1927, May 16, 991
CHAMBERLAIN, C. W., Elected Associate 1926, January 18, 485
CHAMBERLAIN, K. M., Elected Associate 1925, October 26, 483
Charcoal, oxidations on (WRIGHT), 187
Chemical reactions, Apparatus for measuring the velocity of very rapid, II (HARTRIDGE and ROUGHTON), 450
CHITTY, L., Elected Associate 1927, January 17, 988
CHRISTIANSEN, J. A., Note on the velocity of gas-reactions, 438
Circles, illustration of the space representation of (GRACE), 859
CLAPHAM, A. R., Elected Fellow 1926, November 22, 987
CLARK, R. J., A convenient method of distillation of the alkali metals, 953
—— A rapid mercury still, 956
Compton effect in wave mechanics (DIRAC), 500
Condenser battery, over-tension in a, during a sudden discharge (KAPITZA), 144
Congruences, theory of linear (MACMAHON), 337
CONSTABLE, F. H., The behaviour of the centres of activity of saturated surfaces during the initial stages of unimolecular reactions, 172
—— On the stability of copper catalysts produced by thermal decomposition, 432
—— Surface adsorption and the velocity of chemical action at gas-solid interfaces, 593
—— On the effect of continued small additions of poisonous substances on the velocity of gaseous catalytic reaction in closed vessels, 832
COOMBS, H. I., Elected Associate 1926, February 15, 486
Copper catalysts produced by thermal decomposition, stability of (CONSTABLE), 432
Correspondence, the general (m, n) (VAIDYANATHASWAMY), 109
—— the (2, 1) (VAIDYANATHASWAMY), 233
—— the (3, 2) (VAIDYANATHASWAMY), 631
Counter, electrical (EMELÉUS), 85
COX, S. F., Elected Associate 1927, July 25, 992
CRASTER, J. E., Elected Associate 1926, July 26, 488

Digitized by Google

CROWTHER, J. A., A theory of the action of X-rays on living cells, 284
Crystals, developments in the X-ray analysis of single (ASTON), 561
Crystals, tensile deformation of large aluminium, at crystal boundaries (ASTON), 549
Crystals of iron, Hall effect in single (WEBSTER), 800
Cubic surface, projective transformations of a (BURNSIDE), 498
Cubic surface, quadrinodal (TIMMS), 68
Curve, rational normal, simplexes and other configurations upon a (WHITE), 882
Curves, higher singularities of plane algebraic (SCOTT), 206

DEE, P. I., Elected Fellow 1926, November 22, 987
Density, relation between refractive index and (BURNETT), 907
DICKSON, G. H., Elected Associate 1927, July 25, 992
Differential equation of the second order with periodic coefficients, periodic solutions of a linear (INCE), 44
Diffusion pump, theory of the (WERTENSTEIN), 578
DIRAC, P. A. M., Elected Fellow 1926, November 22, 987
—— The adiabatic hypothesis for magnetic fields, 69
—— On quantum algebra, 412
—— The Compton effect in wave mechanics, 500
Distillation of the alkali metals (CLARK), 953
DITCHBURN, R. W., Elected Fellow 1925, November 9, 484
—— The quenching of resonance radiation and the breadths of absorption-lines, 78
—— Notes on Spectrophotometry, 959
Double reflexion (APPLEYARD and SKINNER), 951
Double sixes, on a group of 1440 birational transformations of four variables that arises in considering the projective equivalence of (BURNSIDE), 103
DURRANT, P. J., Elected Fellow 1925, November 23, 484
DYMOND, E. G., Excitation by high velocity electrons, 804

EADIE, G., Elected Associate 1925, October 26, 483
Eddy currents in thin circular cylinders of uniform conductivity due to periodically changing magnetic fields, in two dimensions (CARTER), 901
EDGE, W. L., Elected Fellow 1927, May 16, 991
EDWARDS, B. A., Elected Associate 1925, October 26, 483
ELDER, L. W., Elected Associate 1926, October 25, 986
Electrical counter (EMELÉUS), 85
Electrical doublets, approximation to the motion of two rotating (TAYLOR), 269
Electrode, gas discharges by means of an exploring (EMELÉUS), 531
Electron emissions, differential retarding potential method for the study of the energy distribution of slow (SHARMAN), 922
Electronic emission, secondary, application of the method of the magnetic spectrum to the study of (SHARMAN), 523
Electrons, excitation by high velocity (DYMOND), 804
Elliptic functions, chains of two-two relations and the theory of (BAKER), 92
Elliptic products of Jacobi and the theory of linear congruences (MACMAHON), 337
ELLIS, C. D., and W. A. WOOSTER, The absolute intensities of the γ-rays of radium B and radium C, 717
EMELÉUS, K. G., Notes on the electrical counter, 85
—— Some investigations of gas discharges by means of an exploring electrode, 531
Error of an interpolated value, and on the construction of tables (FISHER and WISHART), 912
Euclid's elements, fifth book of. (Second Addendum to fifth paper.) Proof of Euclid v 19 (HILL), 779
Euclid's postulate of parallels (HILL), 19

FAULKNER, I. J., Elected Associate 1925, October 26, 483
Ferromagnetism, theory of (MAHAJANI), 136
FISHER, R. A., and J. WISHART, On the distribution of the error of an interpolated value, and on the construction of tables, 912
FLETCHER, B. A., Elected Associate 1926, October 25, 986
Fluid, stability of superposed streams of (TAYLOR), 730
Fonctions de Mathieu, sur la non-existence simultanée de deux (MARKOVIĆ), 203
Fourier integral of finite type, summation of a (POLLARD), 373
Fourier integral of finite type, summation of the integral conjugate to the (GRIMSHAW), 871
Fourier integrals and Fourier series,

Digitized by Google

case of distinction between (Grimshaw), 755
Fourier series of a nearly continuous function, summability of the (Hardy and Littlewood), 681
Franck-Cario reactions, influence of an illuminated mercury surface on the (Hirst), 162

Gabbatt, J. P., Note on the pedal locus, 28
—— Note on the extension to higher space of a theorem of Wallace, 361
Gamma ray form by radium E, amount of energy emitted in the (Aston), 935
Gamma rays of radium B and radium C, absolute intensities of the (Ellis and Wooster), 717
Gas discharges by means of an exploring electrode (Emeléus), 531
Gaseous catalytic reaction in closed vessels, velocity of (Constable), 832
Gas reactions, velocity of (Christiansen), 438
Gas-solid interfaces, surface adsorption and the velocity of chemical action at (Constable), 593
Gaunt, J. A., The stopping power of hydrogen atoms for α-particles according to the new quantum theory, 732
—— and W. H. McCrea, The emission of radiation by a quadripole electric moment on the quantum mechanics, 930
Geodetic and dynamical principles, a comparison and connexion (Hargreaves), 493
Gibbs' phase-integral for imperfect gases, evaluation of (Ursell), 685
Gillespie, R. P., Elected Associate 1926, October 25, 986
Goldstein, S., Elected Fellow 1925, November 23, 484
—— The stability of a strut under thrust, when buckling is resisted by a force proportional to the displacement, 120
—— On Mathieu Functions. *See Transactions*, xxiii
Grace, J. H., A theorem of Dr P. Zeeman, 845
—— The pedal planes of a tetrahedron, 853
—— An illustration of the space representation of circles, 859
Green, H. L. H. H., Elected Fellow 1926, February 15, 486
Grimshaw, M. E., A case of distinction between Fourier integrals and Fourier series, 755
—— Summation of the integral conjugate to the Fourier integral of finite type, 871
Group of 1440 birational transformations (Burnside), 103
Group of order 25920 (Burnside), 498

Haldane, J. B. S., A mathematical theory of natural and artificial selection. Part III, 363; Part IV, 607; Part V, 838
—— The change with age of a linkage between Mendelian factors in fowls, 484
—— The comparative genetics of colour in Rodents and Carnivora. *See Biol. Rev.* ii
Hall effect in single crystals of iron (Webster), 800
Hardy, G. H., Note on Ramanujan's arithmetical function $\tau(n)$, 675
—— and J. E. Littlewood, Notes on the theory of series (III): On the summability of the Fourier series of a nearly continuous function, 681
Hare, J. G., Elected Associate 1925, November 23, 484
Hargreaves, R., The problem of relativity in reference to several bodies, 191
—— Relativity in connexion with axial rotation, 198
—— Geodetic and dynamical principles, a comparison and connexion, 493
Hartree, D. R., Some relations between the optical spectra of different atoms of the same electronic structure. II. Aluminium-like and copper-like atoms, 304
Hartridge, H., and F. J. W. Roughton, Improvements in the apparatus for measuring the velocity of very rapid chemical reactions. II, 450
Hawkins, H. L., Some problems in the evolution of Echinoidea. *See Biol. Rev.* ii
Hecht, S., Elected Associate 1925, October 26, 483
Henderson, M. C., Elected Associate 1926, March 15, 487
Herschel's condition and the optical cosine law (Steward), 703
Hickson, S. J., Elected Fellow 1927, January 17, 988
Hill, A. V., Irreversible phenomena in muscular shortening, 486
Hill, M. J. M., On the hypothesis of the obtuse angle, 2
—— On the substitution of Wallis's postulate of similarity for Euclid's postulate of parallels. Addendum, 19
—— On the fifth book of Euclid's elements. (Second Addendum to fifth paper.) Proof of Euclid v 19, 779
Hilton, J., Elected Fellow 1927, January 17, 988.

Digitized by Google

HIRST, H. S., The influence of an illuminated mercury surface on the Franck-Cario reactions, 162
—— Photocatalysis at mercury surfaces, 486
HOLTTUM, R. E., Some problems of plant biology in Malaya. *See Biol. Rev.* II
HOOVER, G. I., Elected Associate 1925, November 23, 484
HUDSON, E. P., Elected Fellow 1926, November 22, 987
HUGHES, W. S., Elected Associate 1925, November 23, 484
HUNTER, W., Lines and planes in a metrical space, 653

INCE, E. L., Periodic solutions of a linear differential equation of the second order with periodic coefficients, 44
—— The second solution of the Mathieu equation, 47
Insular gravity and oceanic isostasy (LARMOR), 130
Integral conjugate to the Fourier integral of finite type (GRIMSHAW), 871
Integral invariants of the affine field (NEWMAN), 262
Interpolated value, and on the construction of tables, distribution of the error of an (FISHER and WISHART), 912
Iron, crystals of, Hall effect (WEBSTER), 800
Isostasy, oceanic (LARMOR), 130
IWATSUKI, T., Elected Associate 1926, October 25, 986

JACKSON, C. H. N., Elected Associate 1925, November 23, 484
Jacobi, elliptic products of, and the theory of linear congruences (MACMAHON), 337
Jacobi's rotating liquid ellipsoid (SEDGWICK), 33
JAMES, C. G. F., On a family of constructs in higher space, 50
—— Some formulae for scrolls and line systems in ordinary space. *See Transactions*, XXIII
—— *Deceased*, 985
JEFFREYS, H., On compressional waves in two superposed layers, 472
—— Wave propagation in strings with continuous and concentrated loads, 768
JENKIN, P. M., Elected Associate 1925, October 26, 483
JESSOP, G., Elected Associate 1926, October 25, 986
JONES, B. M., Wings, 484
JONES, O. A., Elected Associate 1926, October 25, 986

KAPITZA, P., Over-tension in a condenser battery during a sudden discharge, 144
KNIPP, C. T., Elected Associate 1926, October 25, 986

LANDALE, S. E. A., Elected Associate 1927, February 14, 989
LANGLEY, J. N., *Deceased*, 985
LARMOR, J., Insular gravity and oceanic isostasy, 130
—— What determines the resistance and the tilt of an aeroplane? 617
LEWIS, B., Elected Associate 1925, November 23, 484
LEWIS, E. G., Elected Associate 1925, November 23, 484
LEWIS, H. M., Elected Associate 1926, January 18, 485
LEWIS, W. J., *Deceased*, 985
Lines and planes in a metrical space (HUNTER), 653
Lines in threefold space and points of a quadric in fivefold space (WREN), 386
LIPSCHÜTZ, A., On some fundamental laws of ovarian dynamics. *See Biol. Rev.* II
LITTLEWOOD, J. E., and G. H. HARDY, Notes on the theory of series (III): On the summability of the Fourier series of a nearly continuous function, 681
LLOYD, W. V., Elected Associate 1926, October 25, 986
LOB, H., Elected Fellow 1927, February 14, 989
LOWRY, T. M., Recent advances in Stereochemistry, 987

MACCONKEY, C. A., Elected Associate 1926, October 25, 986
MCCREA, W. H., Elected Fellow 1927, January 17, 988
—— The specific heat of carbon dioxide and the form of the CO_2 molecule, 890
—— The specific heat of water vapour and the theory of the dissociation of water vapour at high temperatures, 942
—— and J. A. GAUNT, The emission of radiation by a quadripole electric moment on the quantum mechanics, 930
MACMAHON, P. A., The elliptic products of Jacobi and the theory of linear congruences, 337
MADGWICK, E., The β-ray spectrum of Ra E, 982
—— The absorption and reduction in velocity of β-rays on their passage through matter, 970

Digitized by Google

Magnetic fields, adiabatic hypothesis for (DIRAC), 69
MAHAJANI, G. S., A contribution to the theory of ferromagnetism, 136
MANN, F. G., and W. J. POPE, Metallic complexes with the aliphatic poly-amines, 183
MANTON, S. M., Elected Associate 1926, October 25, 986
MARKOVIĆ, Ž., Sur la non-existence simultanée de deux fonctions de Mathieu, 203
MARSHALL, F. H. A., The conditions governing parturition. *See Biol. Rev.* II
MARTIN, L. H., Some measurements on the absorption of X-rays of long wave-length, 783
MARTIN, N., Elected Associate 1927, July 25, 992
Mathieu equation, the second solution of the (INCE), 47
Mathieu, fonctions de, non-existence simultanée de deux (MARKOVIĆ), 203
Maxwell's law, and the absorption and emission of radiation (MILNE), 465
Mellin's inversion formula (BURKILL), 356
Mercury lines emitted from a discharge tube in a magnetic field, polarisation of (SKINNER), 508
Mercury still, a rapid (CLARK), 956
MEREDITH, C. A., Elected Fellow 1926, March 15, 487
Metallic complexes with the aliphatic poly-amines (MANN and POPE), 183
Metrical space, lines and planes in a (HUNTER), 653
MILLER, A. B., Elected Fellow 1925, November 9, 484
MILNE, E. A., Maxwell's law, and the absorption and emission of radiation, 465
—— The total energy of binding of a heavy atom, 794
MITCHELL, T. C., Elected Associate 1925, October 26, 483
Molecule, specific heat of carbon dioxide and the form of the CO_2 (MCCREA), 890
MORRISON, D. M., Elected Associate 1925, October 26, 483
MOULLIN, E. B., Elected Fellow 1927, January 17, 988
—— On some resistance properties of a certain network containing inductances and capacities and their analogies in a vibrating mechanical system, 391

NATH, V., On the present position of the mitochondria and the Golgi apparatus. *See Biol. Rev.* II
Network, resistance properties of, containing inductances and capacities and their analogies in a vibrating mechanical system (MOULLIN), 391
NEWMAN, M. H. A., Integral invariants of the affine field, 262
NIMMO, R. R., Elected Associate 1926, November 8, 987

Obtuse angle, hypothesis of the (HILL), 2
ONSLOW, H., Elected Associate 1926, March 15, 487
OPPENHEIMER, J. R., Elected Associate 1926, January 18, 485
—— On the quantum theory of vibration-rotation bands, 327
—— On the quantum theory of the problem of the two bodies, 422
Optical cosine law, Herschel's condition and the (STEWARD), 703
Optical illusion due to contrast (THOMSON), 419
Optical spectra of different atoms of the same electronic structure, II (HARTREE), 304
Over-tension in a condenser battery during a sudden discharge (KAPITZA), 144
Oxidations on charcoal (WRIGHT), 187

Pappus' theorem, the figure of (RAO), 496
PARKES, A. S., The mammalian sex-ratio. *See Biol. Rev.* II
Parseval and Riesz-Fischer, generalizations of the theorems of (POLLARD), 865
PATERSON, N. F., Elected Associate 1926, February 15, 486
PEASE, M. S., A note on the inheritance of yellow fat in rabbits, 484
Pedal locus (GABBATT), 28
Pedal planes of a tetrahedron (GRACE), 853
PENTELOW, F. T. K., Elected Fellow 1927, March 14, 990
Pepper, E. D., Elected Associate 1926, July 26, 488
Periodic solutions of a linear differential equation (INCE), 44
Petzval optical condition (STEWARD), 461
Polarizabilities of atomic cores (SWIRLES), 403
Polarization of mercury line (SKINNER), 508
POLLARD, S., The summation of a Fourier integral of finite type, 373
—— On the descriptive form of Taylor's theorem, 383
—— On the generalizations of the theorems of Parseval and Riesz-Fischer, 865
POPE, W. J., and F. G. MANN, Metallic complexes with the aliphatic poly-amines, 183

Digitized by Google

PORTER, E. F., Elected Associate 1926, October 25, 986
Proceedings at the Meetings held during the Session 1925-1926, 482; 1926-1927, 985
PUMPHREY, R. J., Elected Associate 1927, July 25, 992
PUNNETT, R. C., On a case of matriclinous inheritance in *Antirrhinum*, 484
PURVIS, J. E., The absorption spectra of "saturated" and "unsaturated" organic substances, 588

Quadric fourfold in fivefold space (WREN), 386
Quadric, reciprocation of one, into another (BAKER), 22
Quadripole electric moment on the quantum mechanics, emission of radiation by a (GAUNT and MCCREA), 930
Quantum algebra (DIRAC), 412
Quantum mechanics, emission of radiation by a quadripole electric moment on the (GAUNT and MCCREA), 930
Quantum theory of the problem of the two bodies (OPPENHEIMER), 422
Quantum theory of vibration-rotation bands (OPPENHEIMER), 327

Radiation, emission of, by a quadripole electric moment on the quantum mechanics (GAUNT and MCCREA), 930
Radiation, Maxwell's law, and the absorption and emission of (MILNE), 465
Radiation, resonance, and the breadths of absorption-lines (DITCHBURN), 78
Radium B and C, period of decay of (BRACELIN), 150
Radium E, amount of energy emitted in the γ-ray form by (ASTON), 935
Radium E, beta ray spectrum of (MADGWICK), 982
Ramanujan's arithmetical function $\tau(n)$ (HARDY), 675
RAO, C. V. H., On the figure of Pappus' theorem, 496
RATCLIFFE, J. A., Elected Fellow 1926, November 8, 987
—— and M. A. F. BARNETT, On the attenuation of wireless signals in short distance overland transmission, 288
RAWLINS, F. I. G., Elected Fellow 1927, February 14, 989
Reciprocation of one quadric into another (BAKER), 22
REE, A., Elected Associate 1926, February 1, 486
Reflexion, double (APPLEYARD and SKINNER), 951
Refractive index and density, relation between (BURNETT), 907
Relativity in connexion with axial rotation (HARGREAVES), 198
Relativity, problem of, in reference to several bodies (HARGREAVES), 191
RENNIE, J. V. L., Elected Associate 1926, November 8, 987
Resistance properties of a network (MOULLIN), 391
RICHARDS, W. T., Note on the effect of alpha particles on paraffin, 516
Riesz-Fischer, generalizations of the theorems of (POLLARD), 865
ROBINSON, B. W., Elected Fellow 1926, July 26, 488
ROBINSON, M. E., Elected Associate 1925, October 26, 483
Robinson, V. C., Elected Fellow 1925, October 26, 482
Rotating electrical doublets, approximation to the motion of two (TAYLOR), 269
Rotating liquid ellipsoid, Jacobi's (SEDGWICK), 33
Rotating masses of liquid, stability of (BAKER), 1
ROUGHTON, F. J. W., and H. HARTRIDGE, Improvements in the apparatus for measuring the velocity of very rapid chemical reactions, II, 450
RUDGE, W. A. D., Elected Fellow 1926, March 15, 487
RUSSELL, F. S., The vertical distribution of plankton in the sea. *See Biol. Rec.* II

SALAMAN, E., Elected Associate 1926, October 25, 986
SALAMAN, M. H., Elected Fellow 1926, November 22, 987
SAYERS, K. I., Elected Associate 1926, October 25, 986
Scattering, limits of classical (BLACKETT), 698
SCHULMAN, J. H., Elected Associate 1926, October 25, 986
SCOTT, C. A., Elected Associate 1925, December 7, 484
—— On the higher singularities of plane algebraic curves, 206
SEDGWICK, W. F., On the series of forms of Jacobi's rotating liquid ellipsoid, 33
Selection, mathematical theory of natural and artificial (HALDANE), Part III, 363; Part IV, 607; Part V, 838
SETNA, S. B., Elected Associate 1927, May 16, 991
SHARMAN, C. F., The application of the method of the magnetic spectrum to the study of secondary electronic emission, 523
—— A differential retarding potential method for the study of the energy distribution of slow electron emissions, 922

Digitized by Google

Simplexes upon a rational normal curve (WHITE), 882
Singularities of plane algebraic curves, higher (SCOTT), 206
SKINNER, H. W. B., On the polarization of mercury lines emitted from a discharge tube in a magnetic field, 508
—— and E. T. S. APPLEYARD, A case of double reflexion, 951
SMITH, F. P., Elected Fellow 1926, November 8, 987
SMITH, G. F., Elected Associate 1926, October 25, 986
SOMMERFELD, A., Some controversial points in the theory of spectra, 487
Space, higher, family of constructs in (JAMES), 50
Space representation of circles (GRACE), 859
Specific heat of carbon dioxide and the form of the CO_2 molecule (MCCREA), 890
Specific heat of water vapour and the theory of the dissociation of water vapour at high temperatures (MCCREA), 942
Spectrophotometry, notes on (DITCHBURN), 959.
Stability of rotating masses of liquid (BAKER), 1
Stability of a strut under thrust (GOLDSTEIN), 120
Stability of superposed streams of fluid (TAYLOR), 730
STEWARD, G. C., Note on the Petzval optical condition, 461
—— On the addition of the primary aberrations, 584
—— On Herschel's condition and the optical cosine law, 703
—— The aberrations of a symmetrical optical system. *See Transactions*, XXIII
STOCKDALE, D., Elected Fellow 1927, July 25, 991
STRANGEWAYS, D. H., Elected Associate 1927, July 25, 992
Strut, stability of a, under thrust (GOLDSTEIN), 120
SUDA, K., Elected Associate 1926, October 25, 986
SWIRLES, B., Elected Associate 1925, October 26, 483
—— The polarizabilities of atomic cores, 403

Tables, construction of (FISHER and WISHART), 912
TAYLOR, E. M., Elected Associate 1925, October 26, 483
TAYLOR, G. I., An experiment on the stability of superposed streams of fluid, 730
TAYLOR, P. A., An approximation to the motion of two rotating electrical doublets in a plane, 269
TAYLOR, W., Elected Associate 1926, October 25, 986
Taylor's theorem, descriptive form of (POLLARD), 383
Tensile deformation of large aluminium crystals at crystal boundaries (ASTON), 549
Tetrahedron, pedal planes of a (GRACE), 853
THOMAS, L. H., Elected Fellow 1926, November 22, 987
—— The calculation of atomic fields, 542
—— The effect of the orbital velocity of the electrons in heavy atoms on their stopping of α-particles, 713
—— The production of characteristic X-rays by electronic impact, 829
THOMSON, G. P., An optical illusion due to contrast, 419
TIMMS, G., The equianharmonic and harmonic envelopes of a quadrinodal cubic surface, 68
TUNNICLIFFE, H. E., Glutathione. *See Biol. Rev.* II
TURNBULL, H. W., and A. YOUNG, The linear invariants of ten quaternary quadrics. *See Transactions*, XXIII
TURVEY, B. H. C., Elected Associate 1927, January 17, 988
Two-two relations, chains of (BAKER), 92

Unimolecular reactions, activity of saturated surfaces during the initial stages of (CONSTABLE), 172
URSELL, H. D., The evaluation of Gibbs' phase-integral for imperfect gases, 685

VAIDYANATHASWAMY, R., The general (m, n) correspondence, 109
—— the (2, 1) correspondence, 233
—— the pedal (3, 2) correspondence, 631
Valve amplifier for ionisation currents (WYNN-WILLIAMS), 811
VEDY, L. G., Elected Associate 1925, October 26, 483
VEITCH, E. B., Elected Associate 1926, March 15, 487
Velocity of gas-reactions (CHRISTIANSEN), 438
Velocity of very rapid chemical reactions, apparatus for measuring the, II (HARTRIDGE and ROUGHTON), 450
VERNON, M. A., Elected Fellow 1926, November 22, 987
Vibration-rotation bands, quantum theory of (OPPENHEIMER), 327

Digitized by Google

Wallace theorem, extension to higher space of (GABBATT), 361
Wallis's postulate of similarity for Euclid's postulate of parallels (HILL), 19
WANSBROUGH-JONES, O. H., Elected Fellow 1926, November 22, 987
WARD, A. F. H., Elected Fellow 1926, November 8, 987
Water vapour and the theory of the dissociation of water vapour at high temperatures, specific heat of (McCREA), 942
WATKINS, A. E., Elected Fellow 1926, March 1, 487
WATSON, E. E., Elected Associate 1926, October 25, 986
WATSON, S. W., Elected Fellow 1926, March 1, 487
Wave mechanics, Compton effect in (DIRAC), 500
Wave propagation in strings with continuous and concentrated loads (JEFFREYS), 768
Waves, compressional, in two superposed layers (JEFFREYS), 472
WEBSTER, W. L., The Hall effect in single crystals of iron, 800
WERTENSTEIN, L., Elected Associate 1925, October 26, 483
—— Contribution to the theory of the diffusion pump, 578
WHITE, F. P., Simplexes and other configurations upon a rational normal curve, 882
WILKINS, T. R., Elected Associate 1925, October 26, 483
WILSON, A. H., Elected Fellow 1926, November 22, 987
Wireless signals in short distance overland transmission, attenuation of (RATCLIFFE and BARNETT), 288
Wireless transmission, the diurnal variation of ultra-short wave (APPLETON), 155
WISHART, J., and R. A. FISHER, On the distribution of the error of an interpolated value, and on the construction of tables, 912
WOOD, T. B., Animal Calorimetry, 988.
WOODWARD, W. E., Elected Fellow 1926, July 26, 488
WOOSTER, W. A., and C. D. ELLIS, The absolute intensities of the γ-rays of radium B and radium C, 717
WORMELL, T. W., Elected Fellow 1927, February 14, 989
WREN, T. L., The correspondence between lines in threefold space and points of a quadric fourfold in fivefold space, established by a geometrical construction, 386
WRIGHT, W. M., Oxidations on charcoal, 187
WRIGHT, W. R., Elected Associate 1925, October 26, 483
WYNN-WILLIAMS, C. E., Elected Associate 1925, October 26, 483
—— A valve amplifier for ionisation currents, 811

X-ray analysis of single-crystals, developments in the (ASTON), 561
X-rays, action of, on living cells (CROWTHER), 284
X-rays of long wave-length, measurements on the absorption of (MARTIN), 783
X-rays, production of characteristic, by electronic impact (THOMAS), 829

YEATES, J. S., Some problems in the comparison of chromosomes. *See Biol. Rev.* II
YOUNG, A., and H. W. TURNBULL, The linear invariants of ten quaternary quadrics. *See Transactions* XXIII

Zeeman, P., theorem of (GRACE), 845

CAMBRIDGE: PRINTED BY W. LEWIS, M.A., AT THE UNIVERSITY PRESS.

Digitized by Google

Oxford Library
NOV 14 1927
Physics

PROCEEDINGS

OF THE

CAMBRIDGE PHILOSOPHICAL SOCIETY

VOL. XXIII

PART VIII

(WITH TITLE-PAGE, TABLE OF CONTENTS, AND INDEX TO VOL. XXIII)

CAMBRIDGE
AT THE UNIVERSITY PRESS
AND SOLD BY
DEIGHTON, BELL & CO., LIMITED,
AND BOWES & BOWES, CAMBRIDGE
CAMBRIDGE UNIVERSITY PRESS,
FETTER LANE, LONDON, E.C. 4
1927

Price Seven Shillings and Sixpence Net

PRINTED IN GREAT BRITAIN

October 1927.

Digitized by Google

NOTICES.

1. As the stock of early Parts of the *Transactions* and *Proceedings* of the Society is very unequal, the Council is obliged to require that no separate Parts of Volumes I—XI of the *Transactions*, or of Volumes I—XVI of the *Proceedings*, shall be sold without the authority of the Secretaries; to whom also enquiries concerning complete sets should be made.

Members of the Society having early Parts of the publications to dispose of, to add to the sets in the hands of the Society, are invited to communicate with the Secretaries.

2. Letters, and Communications to the Society, should be addressed to one of the Secretaries,

Mr F. P. WHITE, St John's College. [Mathematical]

Mr R. H. FOWLER, Trinity College. [Physical]

Mr F. T. BROOKS, Emmanuel College. [Biological]

Presents for the Library of the Society should be addressed to *The Philosophical Library, New Museums, Cambridge.*

3. Fellows and Honorary Members who desire the *Transactions* of the Society to be sent to them should communicate with the Secretaries if they do not receive the Numbers as they appear.

4. Authors of papers are informed that the Illustrations and Diagrams are executed as far as possible by photographic "process" work; drawings should be on a large scale and on smooth white Bristol board in Indian ink.

Authors may receive, on request, twenty-five offprints of a paper in the *Proceedings* or *Transactions*, without charge. Additional copies may be had by payment to the University Press.

5. *Members of the Society are particularly requested to inform the Secretaries of any change of postal address.*

Digitized by Google

Transactions of the Cambridge Philosophical Society

RECENT PAPERS

Vol. XXII.

No. XXVI. **The Escape of Molecules from an Atmosphere, with special reference to the Boundary of a Gaseous Star.** By E. A. MILNE, M.A., Trinity College, Cambridge. Pp. 35. Price 5/-.

„ XXVII. **Some Problems of Diophantine Approximation: The Analytical Properties of Certain Dirichlet's Series Associated with the Distribution of Numbers to Modulus Unity.** By G. H. HARDY, M.A., Savilian Professor, University of Oxford, and J. E. LITTLEWOOD, M.A., Trinity College, Cambridge. Pp. 16. Price 3/6.

„ XXVIII. **Free Paths in a Non-uniform Rarefied Gas with an Application to the Escape of Molecules from Isothermal Atmospheres.** By J. E. JONES, 1851 Exhibition Senior Research Student, Trinity College, Cambridge. Pp. 22. With Title-page and Index to Vol. XXII. Price 5/6.

Vol. XXIII.

No. I. **On Approximate Continuity.** By M. H. A. NEWMAN, B.A., St John's College, Cambridge. Pp. 18. Price 3/6.

„ II. **A Mathematical Theory of Natural and Artificial Selection.** By J. B. S. HALDANE, M.A., Trinity College, Cambridge. Pp. 23. Price 4/-.

„ III. **On the Form of the Solution of the Equations of Dynamics.** By T. M. CHERRY, B.A., Trinity College, Isaac Newton Student. Pp. 28. Price 4/-.

„ IV. **On the Motion of Spheres, Circular and Elliptic Cylinders through Viscous Liquid.** By W. J. HARRISON, M.A., Clare College, Cambridge. Pp. 18. Price 3/-.

„ V. **Researches in the Theory of Determinants.** By Major P. A. MACMAHON, R.A., F.R.S. Pp. 47. 7/-.

„ VI. **The Optical Constants of Matter.** By C. G. DARWIN, M.A., F.R.S., Christ's College, Tait Professor of Natural Philosophy in the University of Edinburgh. Pp. 31. Price 5/-.

„ VII. **Some Examples of Trajectories Defined by Differential Equations of a Generalised Dynamical Type.** By T. M. CHERRY, PH.D., Fellow of Trinity College, Cambridge. Pp. 32. Price 5/-.

„ VIII. **Some Formulae for Scrolls and Line Systems in Ordinary Space.** By C. G. F. JAMES, Trinity College, Cambridge. Pp. 34. Price 6/-.

„ IX. **The Aberrations of a Symmetrical Optical System.** By G. C. STEWARD, M.A., Fellow of Emmanuel College, Cambridge. Pp. 29. Price 5/-.

„ X. **The Linear Invariants of Ten Quaternary Quadrics.** By H. W. TURNBULL, M.A., Trinity College, Professor of Mathematics in the University of St Andrews, and ALFRED YOUNG, Sc.D., sometime Fellow of Clare College. Pp. 39. Price 6/-.

„ XI. **Mathieu Functions.** By S. GOLDSTEIN, B.A., St John's College. Pp. 34. Price 6/-.

Digitized by Google

CONTENTS.

PAGE

A theorem of Dr P. Zeeman. By Mr J. H. GRACE, Peterhouse . . 845

The pedal planes of a tetrahedron. By Mr J. H. GRACE, Peterhouse . 853

An illustration of the space representation of circles. By Mr J. H. GRACE, Peterhouse 859

On the generalizations of the theorems of Parseval and Riesz-Fischer. By Mr S. POLLARD, Trinity College 865

Summation of the integral conjugate to the Fourier integral of finite type. By MARGARET ELEANOR GRIMSHAW, Newnham College. (Communicated by Mr S. POLLARD) 871

Simplexes and other configurations upon a rational normal curve. By Mr F. P. WHITE, St John's College 882

The specific heat of carbon dioxide and the form of the CO_2 molecule. By W. H. M^{c}CREA, B.A., Trinity College. (Communicated by Mr R. H. FOWLER.) (With Two Text-figures) 890

Eddy currents in thin circular cylinders of uniform conductivity, due to periodically changing magnetic fields, in two dimensions. By Dr F. W. CARTER, St John's College. (Communicated by Mr F. P. WHITE) 901

The relation between refractive index and density. By D. BURNETT, B.A., Clare College 907

On the distribution of the error of an interpolated value, and on the construction of tables. By Dr R. A. FISHER, Gonville and Caius College, and Mr J. WISHART 912

A differential retarding potential method for the study of the energy distribution of slow electron emissions. By C. F. SHARMAN, B.A., King's College. (Communicated by Prof. Sir E. RUTHERFORD.) (With Five Text-figures) 922

The emission of radiation by a quadripole electric moment on the quantum mechanics. By J. A. GAUNT, B.A., Trinity College, and W. H. M^{c}CREA, B.A., Trinity College 930

The amount of energy emitted in the γ-ray form by radium E. By G. H. ASTON, B.A., Trinity Hall. (Communicated by Prof. Sir E. RUTHERFORD.) (With One Text-figure) 935

The specific heat of water vapour and the theory of the dissociation of water vapour at high temperatures. By W. H. M^{c}CREA, B.A., Trinity College. (With One Text-figure) 942

A case of double reflexion. By E. T. S. APPLEYARD, B.A., King's College, and Dr H. W. B. SKINNER, Trinity College. (With One Text-figure) 951

A convenient method of distillation of the alkali metals. By R. J. CLARK, Lecturer in Physics and Carnegie Teaching Fellow in the University of Edinburgh. (Communicated by Prof. Sir E. RUTHERFORD.) (With One Text-figure) 953

A rapid mercury still. By R. J. CLARK, Lecturer in Physics and Carnegie Teaching Fellow in the University of Edinburgh. (Communicated by Prof. Sir E. RUTHERFORD.) (With Three Text-figures) 956

Notes on spectrophotometry. By R. W. DITCHBURN, B.A., Trinity College, Isaac Newton Student. (With Seven Text-figures) 959

The absorption and reduction in velocity of β-rays on their passage through matter. By Dr E. MADGWICK, M.C., Emmanuel College. (Communicated by Prof. Sir E. RUTHERFORD.) (With Nine Text-figures) 970

The β-ray spectrum of Ra E. By Dr E. MADGWICK, M.C., Emmanuel College. (Communicated by Prof. Sir E. RUTHERFORD.) (With One Text-figure) 982

Proceedings at the meetings held during the session 1926—1927 . . 985

Statement of Accounts of the Cambridge Philosophical Society, from January 1, 1926, to December 31, 1926 994

Index to Vol. XXIII 996

Digitized by Google

MAR 27 1926
Physics
131. Sh.

PROCEEDINGS

OF THE

CAMBRIDGE PHILOSOPHICAL SOCIETY

VOL. XXIII

PART I

(WITH PLATE I AND TITLE-PAGE, TABLE OF CONTENTS, AND INDEX TO VOL. XXII)

CAMBRIDGE
AT THE UNIVERSITY PRESS
AND SOLD BY
DEIGHTON, BELL & CO., LIMITED,
AND BOWES & BOWES, CAMBRIDGE
CAMBRIDGE UNIVERSITY PRESS,
FETTER LANE, LONDON, E.C. 4
1926

Price Seven Shillings and Sixpence Net

PRINTED IN GREAT BRITAIN

February 1926.

Digitized by Google

NOTICES.

1. As the stock of early Parts of the *Transactions* and *Proceedings* of the Society is very unequal, the Council is obliged to require that no separate Parts of Volumes I—XI of the *Transactions*, or of Volumes I—XVI of the *Proceedings*, shall be sold without the authority of the Secretaries; to whom also enquiries concerning complete sets should be made.

Members of the Society having early Parts of the publications to dispose of, to add to the sets in the hands of the Society, are invited to communicate with the Secretaries.

2. Letters, and Communications to the Society, should be addressed to one of the Secretaries,

Mr J. GRAY, King's College. [Biological]

Mr F. P. WHITE, St John's College. [Mathematical]

Mr R. H. FOWLER, Trinity College. [Physical]

Presents for the Library of the Society should be addressed to *The Philosophical Library, New Museums, Cambridge.*

3. Fellows and Honorary Members who desire the *Transactions* of the Society to be sent to them should communicate with the Secretaries if they do not receive the Numbers as they appear.

4. Authors of papers are informed that the Illustrations and Diagrams are executed as far as possible by photographic "process" work; drawings should be on a large scale and on smooth white Bristol board in Indian ink.

Authors may receive, on request, twenty-five offprints of a paper in the *Proceedings* or *Transactions*, without charge. Additional copies may be had by payment to the University Press.

5. *Members of the Society are particularly requested to inform the Secretaries of any change of postal address.*

Digitized by Google

Transactions of the Cambridge Philosophical Society

RECENT PAPERS

Vol. XXII.

No. XXIV. **On the Fifth Book of Euclid's Elements (Fifth Paper).** By M. J. M. Hill, Sc.D., LL.D., F.R.S., Astor Professor of Mathematics in the University of London. Pp. 14. Price 2/6.

„ XXV. **The Influence of Electrically Conducting Material within the Earth on Various Phenomena of Terrestrial Magnetism.** By S. Chapman, M.A., Trinity College, and T. T. Whitehead. Pp. 20. Price 3/6.

„ XXVI. **The Escape of Molecules from an Atmosphere, with special reference to the Boundary of a Gaseous Star.** By E. A. Milne, M.A., Trinity College, Cambridge. Pp. 35. Price 5/-.

„ XXVII. **Some Problems of Diophantine Approximation: The Analytical Properties of Certain Dirichlet's Series Associated with the Distribution of Numbers to Modulus Unity.** By G. H. Hardy, M.A., Savilian Professor, University of Oxford, and J. E. Littlewood, M.A., Trinity College, Cambridge. Pp. 16. Price 3/6.

„ XXVIII. **Free Paths in a Non-uniform Rarefied Gas with an Application to the Escape of Molecules from Isothermal Atmospheres.** By J. E. Jones, 1851 Exhibition Senior Research Student, Trinity College, Cambridge. Pp. 22. With Title-page and Index to Vol. xxii. Price 5/6.

Vol. XXIII.

No. I. **On Approximate Continuity.** By M. H. A. Newman, B.A., St John's College, Cambridge. Pp. 18. Price 3/6.

„ II. **A Mathematical Theory of Natural and Artificial Selection.** By J. B. S. Haldane, M.A., Trinity College, Cambridge. Pp. 23. Price 4/-.

„ III. **On the Form of the Solution of the Equations of Dynamics.** By T. M. Cherry, B.A., Trinity College, Isaac Newton Student. Pp. 28. Price 4/-.

„ IV. **On the Motion of Spheres, Circular and Elliptic Cylinders through Viscous Liquid.** By W. J. Harrison, M.A., Clare College, Cambridge. Pp. 18. Price 3/-.

„ V. **Researches in the Theory of Determinants.** By Major P. A. MacMahon, R.A., F.R.S. Pp. 47. 7/-.

„ VI. **The Optical Constants of Matter.** By C. G. Darwin, M.A., F.R.S., Christ's College, Tait Professor of Natural Philosophy in the University of Edinburgh. Pp. 31. Price 5/-.

„ VII. **Some Examples of Trajectories Defined by Differential Equations of a Generalised Dynamical Type.** By T. M. Cherry, Ph.D., Fellow of Trinity College, Cambridge. Pp. 32. Price 5/-.

„ VIII. **Some Formulae for Scrolls and Line Systems in Ordinary Space.** By C. G. F. James, Trinity College, Cambridge. pp. 34. Price 6/-.

Digitized by Google

CONTENTS.

PAGE

The stability of rotating masses of liquid; note to a former paper. By Professor H. F. BAKER 1

On the hypothesis of the obtuse angle. By Professor M. J. M. HILL, Peterhouse. (With Twenty-Six Text-Figures) 2

On the substitution of Wallis's postulate of similarity for Euclid's postulate of parallels. Addendum. By Professor M. J. M. HILL, Peterhouse. (With Three Text-Figures) 19

The reciprocation of one quadric into another. By Professor H. F. BAKER 22

Note on the pedal locus. By Mr J. P. GABBATT, Peterhouse . . . 28

On the series of forms of Jacobi's rotating liquid ellipsoid. By W. F. SEDGWICK, M.A., Trinity College 33

Periodic solutions of a linear differential equation of the second order with periodic coefficients. By Mr E. L. INCE, Trinity College . . 44

The second solution of the Mathieu equation. By Mr E. L. INCE, Trinity College 47

On a family of constructs in higher space. By C. G. F. JAMES, B.A., Trinity College. (With Two Text-Figures) 50

The equianharmonic and harmonic envelopes of a quadrinodal cubic surface. By G. TIMMS, St John's College. (Communicated by Mr W. P. MILNE) 68

The adiabatic hypothesis for magnetic fields. By P. A. M. DIRAC, St John's College. (Communicated by Mr R. H. FOWLER) . . 69

A photographic method of determining the mobility of recoil atoms. By G. H. BRIGGS. (Communicated by Dr J. CHADWICK.) (With One Text-Figure and Plate I) 73

The quenching of resonance radiation and the breadths of absorption-lines. By R. W. DITCHBURN, B.A., Trinity College 78

Notes on the electrical counter. By K. G. EMELÉUS, B.A., St John's College. (Communicated by Mr E. V. APPLETON.) (With Three Text-Figures) 85

On chains of two-two relations and the theory of elliptic functions. By Professor H. F. BAKER 92

Digitized by Google

Digitized by Google

Digitized by Google

Digitized by Google

Digitized by Google